



Nonzero-Sum Differential Games

2

Tamer Başar, Alain Haurie, and Georges Zaccour

Contents

1	Introduction	62
2	A General Framework for m -Player Differential Games	63
2.1	A System Controlled by m Players	63
2.2	Information Structures and Strategies	65
3	Nash Equilibria	68
3.1	Open-Loop Nash Equilibrium (OLNE)	69
3.2	State-Feedback Nash Equilibrium (SFNE)	70
3.3	Necessary Conditions for a Nash Equilibrium	71
3.4	Constructing an SFNE Using a Sufficient Maximum Principle	74
3.5	Constructing an SFNE Using Hamilton-Jacobi-Bellman Equations	75
3.6	The Infinite-Horizon Case	76
3.7	Examples of Construction of Nash Equilibria	77
3.8	Linear-Quadratic Differential Games (LQDGs)	80
4	Stackelberg Equilibria	91
4.1	Open-Loop Stackelberg Equilibria (OLSE)	91
4.2	Feedback Stackelberg Equilibria (FSE)	94
4.3	An Example: Construction of Stackelberg Equilibria	97
4.4	Time Consistency of Stackelberg Equilibria	101
5	Memory Strategies and Collusive Equilibria	102
5.1	Implementing Memory Strategies in Differential Games	102
5.2	An Example	105
6	Conclusion	108
	References	109

T. Başar (✉)

Coordinated Science Laboratory and Department of Electrical and Computer Engineering,
University of Illinois at Urbana-Champaign, Urbana, IL, USA
e-mail: basar1@illinois.edu

A. Haurie

ORDECSYS and University of Geneva, Geneva, Switzerland
GERAD-HEC Montréal PQ, Montreal, QC, Canada
e-mail: ahaurie@gmail.com

G. Zaccour

Department of Decision Sciences, GERAD, HEC Montréal, Montreal, QC, Canada
e-mail: georges.zaccour@gerad.ca

Abstract

This chapter provides an overview of the theory of nonzero-sum differential games, describing the general framework for their formulation, the importance of information structures, and noncooperative solution concepts. Several special structures of such games are identified, which lead to closed-form solutions.

Keywords

Closed-loop information structure · Information structures · Linear-quadratic games · Nash equilibrium · Noncooperative differential games · Non-Markovian equilibrium · Open-loop information structure · State-feedback information structure · Stackelberg equilibrium

1 Introduction

Differential games are games played by agents, also called *players*, who jointly control (through their actions over time, as inputs) a dynamical system described by differential state equations. Hence, the game evolves over a continuous-time horizon (with the length of the horizon known to all players, as common knowledge), and over this horizon each player is interested in optimizing a particular objective function (generally different for different players) which depends on the state variable describing the evolution of the game, on the self-player's action variable, and also possibly on other players' action variables. The objective function for each player could be a reward (or payoff, or utility) function, in which case the player is a maximizer, or it could be a cost (or loss) function, in which case the player would be a minimizer. In this chapter we adopt the former, and this clearly brings in no loss of generality, since optimizing the negative of a reward function would make the corresponding player a minimizer. The players determine their actions in a way to optimize their objective functions, by also utilizing the information they acquire on the state and other players' actions as the game evolves, that is, their actions are generated as a result of the *control policies* they design as mappings from their information sets to their action sets. If there are only two players and their objective functions add up to zero, then this captures the scenario of two totally conflicting objectives – what one player wants to minimize the other one wants to maximize. Such differential games are known as *zero-sum differential games*. Otherwise, a differential game is known to be *nonzero-sum*.

The study of differential games (more precisely, zero-sum differential games) was initiated by Rufus Isaacs at the Rand Corporation through a series of memoranda in the 1950s and early 1960s of the last century. His book Isaacs (1965), published in 1965, after a long delay due to classification of the material it covered, is still considered as the starting point of the field. The early books following Isaacs, such as those by Blaquièrre et al. (1969), Friedman (1971), and Krassovski and Subbotin (1977), all dealt (in most part) with two-player zero-sum differential games. Indeed, initially the focal point of differential games research stayed within the zero-sum domain and was driven by military applications and the presence

of antagonistic elements. The topic of two-player zero-sum differential games is covered in some detail in *this chapter (TPZSDG)* of this *Handbook*.

Motivated and driven by applications in management science, operations research, engineering, and economics (see, e.g., Sethi and Thompson 1981), the theory of differential games was then extended to the case of many players controlling a dynamical system while playing a nonzero-sum game. It soon became clear that nonzero-sum differential games present a much richer set of features than zero-sum differential games, particularly with regard to the interplay between information structures and nature of equilibria. Perhaps the very first paper on this topic, by Case, appeared in 1969, followed closely by a two-part paper by Starr and Ho (1969a,b). This was followed by the publication of a number of books on the topic, by Leitmann (1974), and by Başar and Olsder (1999), with the first edition dating back to 1982, Mehlmann (1988), and Dockner et al. (2000), which focuses on applications of differential games in economics and management science. Other selected key book references are the ones by Engwerda (2005), which is specialized to linear-quadratic differential (as well as multistage) games, Jørgensen and Zaccour (2004), which deals with applications of differential games in marketing, and Yeung and Petrosjan (2005), which focuses on cooperative differential games.

This chapter is on noncooperative nonzero-sum differential games, presenting the basics of the theory, illustrated by examples. It is based in most part on material in Chaps. 6 and 7 of Başar and Olsder (1999) and Chap. 7 of Haurie et al. (2012).

2 A General Framework for m -Player Differential Games

2.1 A System Controlled by m Players

2.1.1 System Dynamics

Consider an n -dimensional dynamical system controlled by a set of m players over a time interval $[t^0, T]$, where $T > t^0$ is a final time that can either be a given data or defined endogenously as the time of reaching a given target, as to be detailed below. For future use, let $M = \{1, \dots, m\}$ denote the players set, that is, the set of all players. This dynamical system has the following elements:

1. A state variable $x \in X \subset \mathbb{R}^n$, and for each player $j \in M$, a control vector $u_j \in U_j \subset \mathbb{R}^{p_j}$, where X and U_j 's are open domains.
2. A state equation (which is an n -dimensional ordinary differential equation) and an initial value for the state (at time t^0)

$$\dot{x}(t) = f(x(t), u(t), t), \quad (2.1)$$

$$x(t^0) = x^0, \quad (2.2)$$

where $\{x(t) : t \in [t^0, T]\}$ is the state trajectory and $\{u(t) \triangleq (u_1(t), \dots, u_m(t)) : t \in [0, T]\}$ is the control (or action) schedule (or simply the control) chosen by

the m players, with $u_j(\cdot)$ generated by player j as a p_j -dimensional function.

Here $\dot{x}(t)$ denotes the time derivative $\frac{d}{dt}x(t)$. The function $f(\cdot, \cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^{p_1+\dots+p_m} \times \mathbb{R} \mapsto \mathbb{R}^n$ is assumed to be continuously differentiable (of class C^1) in x , u , and t .

3. If the control vector generated by the m players is a measurable function of t , or more simply a piecewise continuous function, there is a unique state trajectory solution of (2.1) and (2.2), and each player $j \in M$ receives a cumulative reward over the time horizon $[t^0, T]$:

$$J_j(u(\cdot); x^0, t^0) = \int_{t^0}^T g_j(x(t), u(t), t) dt + S_j(x(T), T), \quad (2.3)$$

where g_j is player j 's instantaneous reward rate and S_j is the terminal reward, also called *salvage value function*. The functions $g_j(\cdot, \cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^{p_1+\dots+p_m} \times \mathbb{R} \mapsto \mathbb{R}$, $j \in M$, are assumed to be continuously differentiable in x , u , and t , and $S_j(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}$, $j \in M$, are assumed to be continuously differentiable in x and t .

2.1.2 Control Constraints

The choice of a control by player j is subject to a pointwise constraint for each $t \in [t^0, T]$

$$u_j(t) \in U_j, \quad t \in [t^0, T], \quad (2.4)$$

where U_j is referred to as the player's *admissible pointwise control set*. In a more general setting, the admissible control set may depend on time t and state $x(t)$. Then, the choice of a control is subject to a constraint

$$u_j(t) \in U_j(x(t), t), \quad t \in [t^0, T], \quad (2.5)$$

where the correspondence, or point-to-set mapping $\{U_j(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R} \mapsto 2^{\mathbb{R}^{p_j}}\}$ is assumed to be *upper-semicontinuous*. In such a case, player j will of course also have to have access to the current value of the state, which brings in the question of what information a player has to have access to before constructing her control; this is related to the information structure of the differential game, without which the formulation of a differential game would not be complete. Information structures will be introduced shortly, in the next subsection.

2.1.3 Target

The determination of the terminal time T can be either prespecified (as part of the initial data), $T \in \mathbb{R}^+$, or the result of the state trajectory reaching a *target*. The target is defined by a *surface* or *manifold* defined by an equation of the form

$$\Theta(x, t) = 0, \quad (2.6)$$

where $\Theta(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}$ is continuously differentiable. The trajectory ends (reaches the target), and the rewards are computed, at the first time T when the condition $\Theta(x(T), T) = 0$ is satisfied.

2.1.4 Infinite-Horizon Games

In economic and engineering applications, one also considers games where the terminal time T may tend to ∞ . The payoff to player j is then defined as

$$J_j(u(\cdot); x^0, t_0) = \int_0^\infty e^{-\rho_j t} g_j(x(t), u(t)) dt. \quad (2.7)$$

Note that player j 's payoff does not include a terminal reward and the reward rate depends explicitly on the running time t through a discount factor $e^{-\rho_j t}$, where ρ_j is a discount rate satisfying $\rho_j \geq 0$, which could be player dependent. An important issue in an infinite-horizon dynamic optimization problem (one-player version of the problem above) is the fact that when the discount rate ρ_j is set to zero, then the integral payoff (2.7) may not be well defined, as the integral may not converge to a finite value for all feasible control paths $u(\cdot)$, and in some cases for none. In such situations, one has to rely on a different notion of optimality, e.g., **overtaking optimality**, a concept well developed in Carlson et al. (1991). We refer the reader to the next chapter (Chap. 3) for a deeper discussion of this topic.

2.2 Information Structures and Strategies

2.2.1 Open Loop Versus State Feedback

To complete the formulation of a differential game, one has to describe precisely the information available to each player (regarding the state and past actions of other players) when they choose their controls at time t . Let us first focus on two information structures of common use in applications of differential games, namely, *open-loop* and *state-feedback* information structures. Letting $v(t)$ denote the information available to a generic player at time t , we say that the information structure is *open loop* if

$$v(t) = \{x^0, t\},$$

that is, the available information is the current time and the initial state. An information structure is *state feedback* if

$$v(t) = \{x(t), t\},$$

that is, the available information is the current state of the system in addition to the current time. We say that a differential game has open-loop (respectively, state-feedback) information structure if every player in the game has open-loop

(respectively, state-feedback) information. It is of course possible for some players to have open-loop information while others have state-feedback information, but we will see later that such a mixed information structure does not lead to a well-defined differential game unless the players who have access to the current value of the state also have access to the initial value of the state, that is,

$$v(t) = \{x(t), x^0, t\}.$$

Another more general information structure is the one with memory, known as *closed-loop with memory*, where at any time t a player has access to the current value of the state and also recalls all past values, that is,

$$v(t) = \{x(s), s \leq t\}.$$

The first two information structures above (open loop and state feedback) are common in *optimal control theory*, i.e., when the system is controlled by only one player. In *optimal control* of a deterministic system, the two information structures are in a sense equivalent. Typically, an optimal state-feedback control is obtained by “synthesizing” the optimal open-loop controls defined from all possible initial states.¹ It can also be obtained by employing *dynamic programming* or equivalently *Bellman’s optimality principle* (Bellman 1957). The situation is, however, totally different for nonzero-sum differential games. The open-loop and state-feedback information structures generally lead to two very different types of differential games, except for the cases of two-player zero-sum differential games (see ► Chap. 8, “Zero-sum Differential Games” in this *Handbook* and also our brief discussion later in this chapter) and differential games with identical objective functions for the players (known as dynamic teams, which are equivalent to optimal control problems as we are dealing with deterministic systems) – or differential games that are *strategically equivalent*² to zero-sum differential games or dynamic team problems. Now, to understand the source of the difficulty in the nonequivalence of two differential games that differ (only) in their information structures, consider the case when the control sets are state dependent, i.e., $u_j(t) \in U_j(x(t), t)$. In the optimal control case, when the only player who controls the system selects a control schedule, she can compute also the associated unique state trajectory. In fact, selecting a control amounts to selecting a trajectory. So, it may be possible to select jointly the control and the associated trajectory to ensure that at each time t the constraint $u(t) \in U(x(t), t)$ is satisfied; hence, it is possible to envision an open-loop control for such a system. Now, suppose that there is another player involved in controlling the system; let us call them players 1 and 2. When player 1 defines

¹See, e.g., the classical textbook on optimal control by Lee and Markus (1972) for examples of synthesis of state-feedback control laws.

²This property will be discussed later in the chapter; in the context of static games, “strategic equivalence” has been discussed in Chap. 1.

her control schedule, she does not know the control schedule of the other player, unless there has been an exchange of information between the two players and a tacit agreement to coordinate their choices of control. Therefore, player 1, not knowing what player 2 will do, cannot decide in advance if her control at time t will be in the admissible set $U_1(x(t), t)$ or not. Hence, in that case, it is impossible for the players to devise feasible and implementable open-loop controls, whereas this would indeed be possible under the state-feedback information structure. The difference between the two information structures is in fact even more subtle, since even when the admissible control sets are not state dependent, knowing at each instant t what the state $x(t)$ is, or not having access to this information will lead to two different types of *noncooperative games in normal form* as we will see in the coming sections.

2.2.2 Strategies

In game theory one calls *strategy* (or *policy* or *law*) a rule that associates an action to the information available to a player at a position of the game. In a differential game, a strategy γ_j for player j is a function that associates to each possible information $v(t)$ at t , a control value $u_j(t)$ in the admissible control set. Hence, for each information structure we have introduced above, we will have a different class of strategies in the corresponding differential game. We make precise below the classes of strategies corresponding to the first two information structures, namely, open loop and state feedback.

Definition 1. Assuming that the admissible control sets U_j are not state dependent, an **open-loop strategy** γ_j for player j ($j \in M$) selects a control action according to the rule

$$u_j(t) = \gamma_j(x^0, t), \quad \forall x^0, \forall t, j \in M, \quad (2.8)$$

where $\gamma_j(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R} \mapsto U_j$ is a function measurable (or piecewise continuous) in t , for each fixed x^0 . The class of all such strategies for player j is denoted by Γ_j^{OL} or simply by Γ_j .

Definition 2. A **state-feedback strategy** γ_j for player j ($j \in M$) selects a control action according to a state-feedback rule

$$u_j(t) = \gamma_j(x(t), t), \quad j \in M, \quad (2.9)$$

where $\gamma_j(\cdot, \cdot) : (x, t) \in \mathbb{R}^n \times \mathbb{R} \mapsto U_j(x, t)$ is a given function that must satisfy the required regularity conditions imposed on feedback controls.³ The class of all such strategies for player j is denoted by Γ_j^{SF} or simply by Γ_j .

³These are conditions which ensure that when all players' strategies are implemented, then the differential equation (2.1) describing the evolution of the state admits a unique piecewise continuously differentiable solution for each initial condition x^0 ; see, e.g., Başar and Olsder (1999).

Remark 1. In the literature on dynamic/differential games, state-feedback strategy is sometimes called “Markovian,” in contrast to “open loop,” with the argument being that the former implies less “commitment” than the latter. Such an interpretation is misleading on two counts. First, one can actually view both classes of strategies as Markovian, since, at each time t , they exploit only the information received at time t . The strategies do not exploit the **history** of the information received up to time t , which is in fact not available. Second, in both cases, a strategy is a full commitment. Using an open-loop strategy means that the player commits, at the initial time, to a fixed time path for her control, that is, her choice of control at each instant of time is predetermined. When using a state-feedback strategy, a player commits to the use of a well-defined servomechanism to control the system, that is, her reaction to the information concerning the state of the system is predetermined. The main advantages of state-feedback strategies lie elsewhere: (i) state-feedback strategies are essential if one has a stochastic differential game (a differential game where the state dynamics are perturbed by disturbance (or noise) with a stochastic description); in fact, if we view a deterministic differential game as the “limit” of a sequence of stochastic games with vanishing noise, we are left with state-feedback strategies. (ii) State-feedback strategies allow us to introduce the refined equilibrium solution concept of “subgame-perfect Nash equilibrium,” which is a concept much appreciated in economic applications, and will be detailed below.

3 Nash Equilibria

Recall the definition of a *Nash equilibrium* for a game in *normal form* (equivalently, *strategic form*).

Definition 3. With the initial state x^0 fixed, consider a differential game in normal form, defined by a set of m players, $M = \{1, \dots, m\}$, and for each player j ($j \in M$) a strategy set Γ_j and a payoff function

$$\bar{J}_j : \Gamma_1 \times \dots \times \Gamma_j \times \dots \times \Gamma_m \mapsto \mathbb{R}, \quad j \in M.$$

Nash equilibrium is a strategy m -tuple $\gamma^* = (\gamma_1^*, \dots, \gamma_m^*)$, such that for each player j the following holds:

$$\bar{J}_j(\gamma^*) \geq \bar{J}_j([\gamma_j, \gamma_{-j}^*]), \quad \forall \gamma_j \in \Gamma_j, \quad (2.10)$$

where $\gamma_{-j}^* := (\gamma_i^* : i \in M \setminus j)$ and $[\gamma_j, \gamma_{-j}^*]$ is the m -tuple obtained when, in γ^* , γ_j^* is replaced by γ_j . In other words, in Nash equilibrium, for each player j , the strategy γ_j^* is the best reply to the $(m - 1)$ -tuple of strategies γ_{-j}^* chosen by the other players.

Corresponding to the first two information structures we have introduced for differential games, we will now define two different *games in normal form*, leading to two different concepts of Nash equilibrium for nonzero-sum differential games.

3.1 Open-Loop Nash Equilibrium (OLNE)

Assume that the admissible control sets U_j , $j \in M$ are not state dependent. If the players use open-loop strategies (2.8), each γ_j defines a unique control schedule $u_j(\cdot) : [0, T] \mapsto U_j$ for each initial state x^0 . The payoff functions for the normal form game are defined by

$$\bar{J}_j(\gamma) = J_j(u(\cdot); x^0, t^0), \quad j \in M, \quad (2.11)$$

where $J_j(\cdot; \cdot, \cdot)$ is the reward function defined in (2.3). Then, we have the following definition:

Definition 4. The control m -tuple $u^*(\cdot) = (u_1^*(\cdot), \dots, u_m^*(\cdot))$ is an **open-loop Nash equilibrium** (OLNE) at (x^0, t^0) if the following holds:

$$J_j(u^*(\cdot); x^0, t^0) \geq J_j([u_j(\cdot), u_{-j}^*(\cdot)]; x^0, t^0), \quad \forall u_j(\cdot), j \in M,$$

where $u_j(\cdot)$ is any admissible control of player j and $[u_j(\cdot), u_{-j}^*(\cdot)]$ is the m -tuple of controls obtained by replacing the j -th block component in $u^*(\cdot)$ by $u_j(\cdot)$.

Note that in the OLNE, for each player j , $u_j^*(\cdot)$ solves the optimal control problem

$$\max_{u_j(\cdot)} \left\{ \int_{t^0}^T g_j(x(t), [u_j(t), u_{-j}^*(t)], t) dt + S_j(x(T)) \right\},$$

subject to the state equation

$$\dot{x}(t) := \frac{d}{dt}x(t) = f(x(t), [u_j(t), u_{-j}^*(t)], t), \quad x(t^0) = x^0, \quad (2.12)$$

control constraints $u_j(t) \in U_j$, and target $\Theta(\cdot, \cdot)$. Further note that OLNE strategies will in general also depend on the initial state x^0 , but this is information available to each player under the open-loop information structure.

3.2 State-Feedback Nash Equilibrium (SFNE)

Now consider a differential game with the state-feedback information structure. The system is then driven by a state-feedback strategy m -tuple $\gamma(\mathbf{x}, t) = (\gamma_j(\mathbf{x}, t); j \in M)$, with $\gamma_j \in \Gamma_j^{\text{SF}}$ for $j \in M$. Its dynamics are thus defined by

$$\dot{x}(t) := \frac{d}{dt}x(t) = f(x(t), \gamma(x(t), t), t), \quad x(t^0) = x^0. \quad (2.13)$$

The normal form of the game, at (x^0, t^0) , is now defined by the payoff functions⁴

$$\bar{J}_j(\gamma; x^0, t^0) = \int_{t^0}^T g_j(x(t), \gamma(x(t), t), t) dt + S_j(x(T)), \quad (2.14)$$

where, for each fixed x^0 , $x(\cdot) : [t^0, T] \mapsto \mathbb{R}^n$ is the state trajectory solution of (2.13).

In line with the convention in the OL case, let us introduce the notation

$$\gamma_{-j}(t, x(t)) \triangleq (\gamma_1(x(t), t), \dots, \gamma_{j-1}(x(t), t), \gamma_{j+1}(x(t), t), \dots, \gamma_m(x(t), t)),$$

for the *strategy* $(m-1)$ -tuple where the strategy of player j does not appear.

Definition 5. The state-feedback m -tuple $\gamma^* = (\gamma_1^*, \dots, \gamma_m^*)$ is a **state-feedback Nash equilibrium** (SFNE) on⁵ $X \times [t^0, \mathcal{T}]$ if for any initial data $(x^0, t^0) \in X \times [0, \mathcal{T}] \subset \mathbb{R}^n \times \mathbb{R}^+$, the following holds:

$$\bar{J}_j(\gamma^*; x^0, t^0) \geq \bar{J}_j([\gamma_j(\cdot), \gamma_{-j}^*(\cdot)]; x^0, t^0), \quad \forall \gamma_j \in \Gamma_j^{\text{SF}}, \quad j \in M,$$

where $[\gamma_j, \gamma_{-j}^*]$ is the m -vector of strategies obtained by replacing the j -th block component in γ^* by γ_j .

In other words, $\{u_j^*(t) \equiv \gamma_j^*(x^*(t), t) : t \in [t^0, T]\}$, where $x^*(\cdot)$ is the equilibrium trajectory generated by γ^* from (x^0, t^0) , solves the optimal control problem

$$\max_{u_j(\cdot)} \left\{ \int_{t^0}^T g_j \left(x(t), \left[u_j(t), \gamma_{-j}^*(x(t), t) \right], t \right) dt + S_j(x(T)) \right\}, \quad (2.15)$$

⁴With a slight abuse of notation, we have included here also the pair (x^0, t^0) as an argument of \bar{J}_j , since under the SF information γ does not have (x^0, t^0) as an argument for $t > t^0$.

⁵We use \mathcal{T} instead of T because, in a general setting, T may be endogenously defined as the time when the target is reached.

subject to the state equation

$$\dot{x}(t) = f(x(t), [u_j(t), \gamma_{-j}^*(x(t), t)], t), \quad x(t^0) = x^0, \quad (2.16)$$

control constraints $u_j(t) \in U_j(x(t), t)$, and target $\Theta(\cdot, \cdot)$. We can also say that γ_j^* is the optimal state-feedback control $u_j^*(\cdot)$ for the problem (2.15) and (2.16). We also note that the single-player optimization problem (2.15) and (2.16) is a standard optimal control problem whose solution can be expressed in a way compatible with the state-feedback information structure, that is, solely as a function of the current value of the state and current time, and not as a function of the initial state and initial time. The remark below further elaborates on this point.

Remark 2. Whereas an open-loop Nash equilibrium is defined only for the given initial data, here the definition of a state-feedback Nash equilibrium asks for the equilibrium property to hold for all initial points, or data, in a region $X \times [t^0, \mathcal{T}] \subset \mathbb{R}^n \times \mathbb{R}^+$. This is tantamount to asking a state-feedback Nash equilibrium to be **subgame perfect** (Selten 1975), in the parlance of game theory, or **strongly time consistent** (Başar 1989). Indeed, even if the state trajectory is perturbed, either because a player has had a “trembling hand” or an unforeseen small shock happened, holding on to the same state-feedback strategy will still constitute a Nash equilibrium in the limit as the perturbations vanish; this property is more pronounced in the case of linear-quadratic differential games (games where the state dynamics are linear, payoff functions are jointly quadratic in the state and the controls, and the time horizon is fixed), in which case the stochastic perturbations in the state equation do not have to be vanishingly small as long as they have zero mean (Başar 1976, 1977). It should be clear that open-loop Nash equilibrium strategies do not possess such a property.

3.3 Necessary Conditions for a Nash Equilibrium

For the sake of simplicity in the exposition below, we will henceforth restrict the target set to be defined by the simple given of a terminal time, that is, the set $\{(t, x) : t = T\}$. Also the control constraint set U_j , $j \in M$ will be taken to be independent of state and time. As noted earlier, at a Nash equilibrium, each player solves an optimal control problem where the system’s dynamics are influenced by the strategic choices of the other players. We can thus write down necessary optimality conditions for each of the m optimal control problems, which will then constitute a set of necessary conditions for a Nash equilibrium. Throughout, we make the assumption that sufficient regularity holds so that all the derivatives that appear in the necessary conditions below exist.

3.3.1 Necessary Conditions for an OLNE

By using the necessary conditions for an open-loop optimal control, obtained, e.g., from the *maximum principle* (see, e.g., Başar and Olsder 1999; Bryson et al. 1975), we arrive at the conditions (2.17), (2.18), (2.19), (2.20), and (2.21) below, which are necessary for an open-loop Nash equilibrium. Let us introduce the individual Hamiltonians, with H_j being the Hamiltonian for player j ,⁶

$$H_j(x, u, \lambda_j, t) = g_j(x, u, t) + \lambda_j(t)f(x, u, t), \quad (2.17)$$

where $\lambda_j(\cdot)$ is the *adjoint* (or *costate*) variable, which satisfies the *adjoint variational equation* (2.18), along with the *transversality condition* (2.19):

$$\dot{\lambda}_j(t) = -\frac{\partial}{\partial x} H_j|_{x^*(t), u^*(t), t}, \quad (2.18)$$

$$\lambda_j(T) = \frac{\partial}{\partial x} S_j|_{x^*(T), T}. \quad (2.19)$$

Further, H_j is maximized with respect to u_j , with all other players' controls fixed at NE, that is,

$$u_j^*(t) = \arg \max_{u_j \in U_j} H_j(x^*(t), u_j, u_{-j}^*(t), \lambda_j(t), t). \quad (2.20)$$

If the solution to the maximization problem above is in the interior of U_j , then naturally a necessary condition is for the first derivative to vanish at u_j^* for all t , that is,

$$\frac{\partial}{\partial u_j} H_j|_{x^*(t), u^*(t), t} = 0, \quad (2.21)$$

and for the Hessian matrix of second derivatives (with respect to u_j) to be nonnegative definite.

3.3.2 Necessary Conditions for SFNE

The state-feedback NE can be obtained in various different ways. One could again use the approach above, but paying attention to the fact that in the optimal control problem faced by a generic player, the other players' strategies are now dependent on the current value of the state. A second approach would be to adapt to this problem the method used in optimal control to directly obtain state-feedback controls (i.e., dynamic programming). We discuss here both approaches, first in this subsection the former. The Hamiltonian for player j is again:

$$H_j(x, u, \lambda_j, t) = g_j(x, u, t) + \lambda_j(t)f(x, u, t). \quad (2.22)$$

⁶We use the convention that $\lambda_j(t)f(\cdot, \cdot, \cdot)$ is the scalar product of two n dimensional vectors $\lambda_j(t)$ and $f(\cdot, \cdot, \cdot)$.

The controls u_i for $i \in M \setminus j$ are now defined by the state-feedback rules $\gamma_i^*(x, t)$. Along the equilibrium trajectory $\{x^*(t) : t \in [t^0, T]\}$, the optimal control of player j is $u_j^*(t) = \gamma_j^*(x^*(t), t)$. Then, as the counterpart of (2.18) and (2.19), we have $\lambda_j(\cdot)$ satisfying (as a necessary condition)

$$\dot{\lambda}_j(t) = - \left(\frac{\partial}{\partial x} H_j + \sum_{i \in M \setminus j} \frac{\partial}{\partial u_i} H_j \frac{\partial}{\partial x} \gamma_i^* \right) \Big|_{x^*(t), u^*(t), t}, \quad (2.23)$$

$$\lambda_j(T) = \frac{\partial}{\partial x} S_j \Big|_{x^*(T), T}, \quad (2.24)$$

where the second term in (2.23), involving a summation, is a reflection of the fact that H_j depends on x not only through g_j and f but also through the strategies of the other players. The presence of this extra term clearly makes the necessary condition for the state-feedback solution much more complicated than for open-loop solution.

Again, $u_j^*(t) = \gamma_j^*(x^*(t), t)$ maximizes the Hamiltonian H_j for each t , with all other variables fixed at equilibrium:

$$u_j^*(t) = \arg \max_{u_j \in U_j} H_j(x^*(t), u_j, \gamma_{-j}^*(x^*(t), t), \lambda_j, t). \quad (2.25)$$

If the solution to the maximization problem above is in the interior of U_j , then as in (2.21) a necessary condition is for the first derivative to vanish at u_j^* for all t , that is,

$$\frac{\partial}{\partial u_j} H_j \Big|_{x^*(t), u_j^*(t), \gamma_{-j}^*(x^*(t), t), t} = 0, \quad (2.26)$$

and for the Hessian matrix of second derivatives (with respect to u_j) to be nonnegative definite.

Remark 3. The summation term in (2.23) is absent in three important cases: (i) in optimal control problems ($m = 1$), since $\frac{\partial}{\partial u} H \frac{\partial u}{\partial x} = 0$; (ii) in two-person zero-sum differential games, because $H_1 \equiv -H_2$ so that for player 1, $\frac{\partial}{\partial u_2} H_1 \frac{\partial u_2}{\partial x} = -\frac{\partial}{\partial u_2} H_2 \frac{\partial u_2}{\partial x} = 0$, and likewise for player 2; and (iii) in open-loop nonzero-sum differential games, because $\frac{\partial u_j}{\partial x} = 0$. It would also be absent in nonzero-sum differential games with state-feedback information structure that are **strategically equivalent** (Başar and Olsder 1999) to (i) (single objective) team problems (which in turn are equivalent to single-player optimal control problems) or (ii) two-person zero-sum differential games.

3.4 Constructing an SFNE Using a Sufficient Maximum Principle

As alluded to above, the necessary conditions for an SFNE as presented are not very useful to compute a state-feedback Nash equilibrium, as one has to infer the form of the partial derivatives of the equilibrium strategies, in order to write the adjoint equations (2.24). However, as an alternative, the sufficient maximum principle given below can be a useful tool when one has an a priori guess of the class of equilibrium strategies (see, Haurie et al. 2012, page 249).

Theorem 1. *Assume that the terminal reward functions S_j are continuously differentiable and concave, and let $X \subset \mathbb{R}^n$ be a state constraint set where the state $x(t)$ belongs for all t . Suppose that an m -tuple $\gamma^* = (\gamma_1^*, \dots, \gamma_m^*)$ of state-feedback strategies $\gamma_j : X \times [t^0, T] \mapsto \mathbb{R}^{m_j}$, $j \in M$, is such that*

- (i) $\gamma^*(x, t)$ is continuously differentiable in x almost everywhere, and piecewise continuous in t ;
- (ii) $\gamma^*(x, t)$ generates at (x^0, t^0) a unique trajectory $x^*(\cdot) : [t^0, T] \mapsto X$, solution of

$$\dot{x}(t) = f(x(t), \gamma^*(x, t), t), \quad x(t^0) = x^0,$$

which is absolutely continuous and remains in the interior of X ;

- (iii) there exist m **costate vector** functions $\lambda_j(\cdot) : [t^0, T] \mapsto \mathbb{R}^n$, which are absolutely continuous and such that, for all $j \in M$, if we define the Hamiltonians

$$\begin{aligned} H_j(x(t), [u_j, u_{-j}], \lambda_j(t), t) \\ = g_j(x(t), [u_j, u_{-j}], t) + \lambda_j(t) f(x(t), [u_j, u_{-j}], t), \end{aligned}$$

and the equilibrium Hamiltonians

$$\mathcal{H}_j^*(x^*(t), \lambda_j(t), t) = \max_{u_j \in U_j} H_j(x^*(t), [u_j, \gamma_{-j}^*(x^*(t), t)], \lambda_j(t), t), \quad (2.27)$$

the maximum in (2.27) is reached at $\gamma_j^*(x^*(t), t)$, i.e.,

$$\mathcal{H}_j^*(x^*(t), \lambda_j(t), t) = \max_{u_j \in U_j} H_j(x^*(t), \gamma^*(x^*(t), t), \lambda_j(t), t); \quad (2.28)$$

- (iv) the functions $x \mapsto \mathcal{H}_j^*(x, \lambda_j(t), t)$ where \mathcal{H}_j^* is defined as in (2.27), but at position (t, x) , are continuously differentiable and concave for all $t \in [t^0, T]$ and $j \in M$;
- (v) the costate vector functions $\lambda_j(\cdot)$, $j \in M$, satisfy the following adjoint differential equations for almost all $t \in [t^0, T]$,

$$\dot{\lambda}_j(t) = -\frac{\partial}{\partial x} \mathcal{H}_j^*|_{(x^*(t), \lambda_j(t), t)}, \quad (2.29)$$

along with the **transversality conditions**

$$\lambda_j(T) = \frac{\partial}{\partial x} S_j|_{(x^*(T), T)}. \quad (2.30)$$

Then, $(\gamma_1^*, \dots, \gamma_m^*)$ is an SFNE at (x^0, t^0) .

3.5 Constructing an SFNE Using Hamilton-Jacobi-Bellman Equations

We now discuss the alternative dynamic programming approach which delivers the state-feedback solution directly without requiring synthesis or guessing of the solution. The following theorem captures the essence of this effective tool for SFNE (see, Başar and Olsder 1999, page 322; Haurie et al. 2012, page 252).

Theorem 2. *Suppose that there exists an m -tuple $\gamma^* = (\gamma_1^*, \dots, \gamma_m^*)$ of state-feedback laws, such that*

- (i) *for any admissible initial point (x^0, t^0) , there exists a unique, absolutely continuous solution $t \in [t^0, T] \mapsto x^*(t) \in X \subset \mathbb{R}^n$ of the differential equation*

$$\dot{x}^*(t) = f(x^*(t), \gamma_1^*(t, x^*(t)), \dots, \gamma_m^*(t, x^*(t)), t), \quad x^*(t^0) = x^0;$$

- (ii) *there exist continuously differentiable value functionals $V_j^* : X \times [t^0, T] \mapsto \mathbb{R}$, such that the following coupled Hamilton-Jacobi-Bellman (HJB) partial differential equations are satisfied for all $(x, t) \in X \times [t^0, T]$*

$$-\frac{\partial}{\partial t} V_j^*(x, t) = \max_{u_j \in U_j} \left\{ g_j(x, [u_j, \gamma_{-j}^*(x, t)], t) + \frac{\partial}{\partial x} V_j^*(x, t) f(x, [u_j, \gamma_{-j}^*(x, t)], t) \right\} \quad (2.31)$$

$$= g_j(x, [\gamma^*(x, t)], t) + \frac{\partial}{\partial x} V_j^*(x, t) f(x, \gamma^*(x, t), t); \quad (2.32)$$

- (iii) *the boundary conditions*

$$V_j^*(x, T) = S_j(x), \quad (2.33)$$

are satisfied for all $x \in X$ and $j \in M$.

Then, $\gamma_j^*(x, t)$, is a maximizer of the right-hand side of the HJB equation for player j , and the m -tuple $(\gamma_1^*, \dots, \gamma_m^*)$ is an SFNE at every initial point $(x^0, t^0) \in X \times [t^0, T]$.

Remark 4. Note that once a complete set of value functionals, $\{V_j, j \in M\}$, is identified, then (2.31) directly delivers the Nash equilibrium strategies of the players in state-feedback form. Hence, in this approach one does not have to guess the SFNE but rather the structure of each player's value function; this can be done in a number of games, with one such class being linear-quadratic differential games, as we will see shortly. Also note that Theorem 2 provides a set of sufficient conditions for SFNE, and hence once a set of strategies are found satisfying them, we are assured of their SFNE property. Finally, since the approach entails dynamic programming, it directly follows from (2.31) that a natural restriction of the set of SFNE strategies obtained for the original differential game to a shorter interval $[s, T]$, with $s > t^0$, constitutes an SFNE for the differential game which is similarly formulated but on the shorter time interval $[s, T]$. Hence, the SFNE is subgame perfect and strongly time consistent.

3.6 The Infinite-Horizon Case

Theorems 1 and 2 were stated under the assumption that the time horizon is finite. If the planning horizon is infinite, then the transversality or boundary conditions, that is, $\lambda_j(T) = \frac{\partial S_j}{\partial x}(x(T), T)$ in Theorem 1 and $V_j(x(t), T) = S_j(x(t), T)$ in Theorem 2, have to be modified. Below we briefly state the required modifications, and work out a scalar example in the next subsection to illustrate this.

If the time horizon is infinite, the dynamic system is autonomous (i.e., f does not explicitly depend on t) and the objective functional of player j is as in (2.7), then the transversality conditions in Theorem 1 are replaced by the limiting conditions:

$$\lim_{t \rightarrow +\infty} e^{-\rho_j t} q_j(t) = 0, \quad \forall j \in M, \quad (2.34)$$

where $q_j(t) = e^{\rho_j t} \lambda_j(t)$ is the so-called current-value costate variable. In the coupled set of HJB equations of Theorem 2, the value function $V_j^*(x, t)$ is multiplicatively decomposed as

$$V_j^*(x, t) = e^{-\rho_j t} \mathcal{V}_j^*(x), \quad (2.35)$$

and the boundary condition (2.33) is replaced by

$$\mathcal{V}_j^*(x, t) \rightarrow 0, \quad \text{when } t \rightarrow \infty, \quad (2.36)$$

which is automatically satisfied if $\mathcal{V}_j^*(x)$ is bounded.

3.7 Examples of Construction of Nash Equilibria

We consider here a two-player infinite-horizon differential game with scalar linear dynamics and quadratic payoff functions, which will provide an illustration of the results of the previous subsection, also to be viewed as illustration of Theorems 1 and 2 in the infinite-horizon case.

Let $u_j(t)$ be the scalar control variable of player j , $j = 1, 2$, and $x(t)$ be the state variable, with $t \in [0, \infty)$. Let player j 's optimization problem be given by

$$\max_{u_j} \left\{ J_j = \int_0^{\infty} e^{-\rho t} \left(u_j(t) \left(\kappa - \frac{1}{2} u_j(t) \right) - \frac{1}{2} \varphi x^2(t) \right) dt \right\},$$

such that $\dot{x}(t) = u_1(t) + u_2(t) - \alpha x(t)$, $x(0) = x^0$,

where φ and κ are positive parameters, $0 < \alpha < 1$, and $\rho > 0$ is the discount parameter. This game has the following features: (i) the objective functional of player j is quadratic in the control and state variables and only depends on the player's own control variable; (ii) there is no interaction (coupling) either between the control variables of the two players or between the control and the state variables; (iii) the game is fully symmetric across the two players in the state and the control variables; and (iv) by adding the term $e^{-\rho t} (u_i(t) (\kappa - \frac{1}{2} u_i(t)))$, $i \neq j$, to the integrand of J_j , for $j = 1, 2$, we can make the two objective functions identical:

$$J := \int_0^{\infty} e^{-\rho t} \left(u_1(t) \left(\kappa - \frac{1}{2} u_1(t) \right) + u_2(t) \left(\kappa - \frac{1}{2} u_2(t) \right) - \frac{1}{2} \varphi x^2(t) \right) dt. \quad (2.37)$$

The significance of this last feature will become clear shortly when we discuss the OLNE (next). Throughout the analysis below, we suppress the time argument when no ambiguity may arise.

Open-Loop Nash Equilibrium (OLNE). We first discuss the significance of feature (iv) exhibited by this scalar differential game. Note that when the information structure of the differential game is open loop, adding to the objective function of a player (say, player 1) terms that involve only the control of the other player (player 2) does not alter the optimization problem faced by player 1. Hence, whether player j maximizes J_j , or J given by (2.37), makes no difference as far as the OLNE of the game goes. Since this applies to both players, it readily follows that every OLNE of the original differential game is also an OLNE of the single-objective optimization problem (involving maximization of J by each player). In such a case, we say that the two games are *strategically equivalent*, and note that the second game (described by the single-objective functional J) is a dynamic team. Now, Nash equilibrium (NE) in teams corresponds to person-by-optimality, and not team optimality (which means joint optimization by members of the team), but when every person-by-person optimal solution is also team optimal (the reverse implication is always true),

then one can obtain all NE of games strategically equivalent to a particular dynamic team by solving for team optimal (equivalently, globally optimal) solutions of the team. Further, when solving for team-optimal solutions in deterministic teams, whether the information structure is open loop or state feedback does not make any difference, as mentioned earlier. In the particular dynamic team of this example, since J is strictly concave in u_1 and u_2 and x (jointly), and the state equation is linear, every person-by-person optimal solution is indeed team optimal, and because of strict concavity the problem admits a unique globally optimal solution. Hence, the OLNE of the original game exists and is unique.

Having established the correspondence with a deterministic concave team and thereby the existence of a unique OLNE, we now turn to our main goal here, which is to apply the conditions obtained earlier for OLNE to the differential game at hand. Toward that end, we introduce the current-value Hamiltonian of player j :

$$\mathcal{H}_j(x, \lambda, u_1, u_2) = u_j \left(\kappa - \frac{1}{2}u_j \right) - \frac{1}{2}\varphi x^2 + q_j(u_1 + u_2 - \alpha x), \quad i = 1, 2,$$

where $q_j(t)$ is the current-value costate variable, at time t , defined as

$$q_j(t) = e^{\rho_j t} \lambda_j(t). \quad (2.38)$$

Being strictly concave in u_j , \mathcal{H}_j admits a unique maximum, achieved by

$$u_j = \kappa + q_j, \quad j = 1, 2. \quad (2.39)$$

Note that the Hamiltonians of both players are strictly concave in x , and hence the equilibrium Hamiltonians also are. Then, the equilibrium conditions read:

$$\begin{aligned} \dot{q}_j &= \rho q_j - \frac{\partial}{\partial x} \mathcal{H}_j = (\rho + \alpha)q_j + \varphi x, & \lim_{t \rightarrow +\infty} e^{-\rho t} q_j(t) &= 0, \quad j = 1, 2, \\ \dot{x} &= 2\kappa + q_1 + q_2 - \alpha x, & x(0) &= x^0. \end{aligned}$$

It is easy to see that $q_1(t) = q_2(t) =: q(t)$, $\forall t \in [0, \infty)$, and therefore $u_1(t) = u_2(t)$, $\forall t \in [0, \infty)$. This is not surprising given the symmetry of the game. We then have a two-equation differential system in x and q :

$$\begin{pmatrix} \dot{x} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} -\alpha & 2 \\ \varphi & \rho + \alpha \end{pmatrix} \begin{pmatrix} x \\ q \end{pmatrix} + \begin{pmatrix} 2\gamma \\ 0 \end{pmatrix}.$$

We look for the solution of this system converging to the steady state which is given by

$$(x_{ss}, q_{ss}) = \left(\frac{2\kappa(\alpha + \rho)}{\alpha^2 + \alpha\rho + 2\varphi}, -\frac{2\kappa\varphi}{\alpha^2 + \alpha\rho + 2\varphi} \right).$$

The solution can be written as

$$\begin{aligned} x(t) &= (x^0 - x_{ss})e^{\mu_1 t} + x_{ss}, \\ q(t) &= -(x^0 - x_{ss}) \frac{2\varphi}{2\alpha + \rho + \sqrt{(2\alpha + \rho)^2 + 8\varphi}} e^{\mu_1 t} + q_{ss}, \end{aligned}$$

where μ_1 is the negative eigenvalue of the matrix associated with the differential equations system and is given by

$$\mu_1 = \frac{1}{2}(\rho - \sqrt{(2\alpha + \rho)^2 + 8\varphi}).$$

Using the corresponding expression for $q(t)$ for q_j in (2.39) leads to the OLNE strategies (which are symmetric).

State-Feedback Nash Equilibrium (SFNE). The strategic equivalence between the OL differential game and a team problem we established above does not carry over to the differential game with state-feedback information structure, since adding any term to J_1 that involves control u_2 of player 2 will alter the optimization problem faced by player 1, since u_2 depends on u_1 through the state x . Hence, the example system is a genuine game under SF information, and therefore the only way to obtain its SFNE would be to resort to Theorem 2 in view of the extension to the infinite horizon as discussed in Sect. 3.6. The HJB equation for player j , written for the current-value function $\mathcal{V}_j(x) = e^{\rho t} V_j(t, x)$, is

$$\rho \mathcal{V}_j(x) = \max_{u_j} \left[u_j \left(\kappa - \frac{1}{2} u_j \right) - \frac{1}{2} \varphi x^2 + \frac{\partial}{\partial x} \mathcal{V}_j(x) (u_1 + u_2 - \alpha x) \right]. \quad (2.40)$$

Being strictly concave in u_j , the RHS of (2.40) admits a unique maximum, with the maximizing solution being

$$u_j(x) = \kappa + \frac{\partial}{\partial x} \mathcal{V}_j(x). \quad (2.41)$$

Given the symmetric nature of this game, we focus on symmetric equilibrium strategies. Taking into account the linear-quadratic specification of the differential game, we make the informed guess that the current-value function is quadratic (because the game is symmetric, and we focus on symmetric solutions, the value function is the same for both players), given by

$$\mathcal{V}_j(x) = \frac{a}{2} x^2 + bx + c, \quad j = 1, 2,$$

where a, b, c are parameters yet to be determined. Using (2.41) then leads to $u_j(x) = \kappa + ax + b$. Substituting this into the RHS of (2.40), we obtain

$$\frac{1}{2}(3a^2 - 2a\alpha - \varphi)x^2 + (3ab - b\alpha + 2a\kappa)x + \frac{1}{2}(3b^2 + 4b\kappa + \kappa^2).$$

The LHS of (2.40) reads

$$\rho \left(\frac{a}{2}x^2 + bx + c \right),$$

and equating the coefficients of x^2 , x and the constant term, we obtain three equations in the three unknowns, a , b , and c . Solving these equations, we get the following coefficients for the noncooperative value functions:

$$\begin{aligned} a &= \frac{\rho + 2\alpha \pm \sqrt{(\rho + 2\alpha)^2 + 16\varphi}}{6}, \\ b &= \frac{-2a\kappa}{3a - (\rho + \alpha)}, \\ c &= \frac{\kappa^2 + 4b\kappa + 3b^2}{2\rho}. \end{aligned}$$

Remark 5. The coefficient a is the root of a second-degree polynomial having two roots: one positive and one negative. The selection of the negative root

$$a = \frac{\rho + 2\alpha - \sqrt{(\rho + 2\alpha)^2 + 16\varphi}}{6},$$

guarantees the global stability of the state trajectory. The resulting noncooperative equilibrium state trajectory is given by

$$x^*(t) = \left[x^0 + \frac{2(\kappa + b)}{2a - \alpha} \right] e^{(2a - \alpha)t} - \frac{2(\kappa + b)}{2a - \alpha}.$$

The state dynamics of the game has a globally asymptotically stable steady state if $2a - \alpha < 0$. It can be shown that to guarantee this inequality and therefore global asymptotic stability, the only possibility is to choose $a < 0$.

3.8 Linear-Quadratic Differential Games (LQDGs)

We have seen in the previous subsection, within the context of a specific scalar differential game, that linear-quadratic structure (linear dynamics and quadratic payoff functions) enables explicit computation of both OLNE and SFNE strategies (for the infinite-horizon game). We now take this analysis a step further, and discuss the general class of linear-quadratic (LQ) games, but in finite horizon, and show

that the LQ structure leads (using the necessary and sufficient conditions obtained earlier for OLNE and SFNE, respectively) to computationally feasible equilibrium strategies. Toward that end, we first make it precise in the definition that follows the class of LQ differential games under consideration (we in fact define a slightly larger class of DGs, namely, affine-quadratic DGs, where the state dynamics are driven by also a known exogenous input). Following the definition, we discuss characterization of the OLNE and SFNE strategies, in that order. Throughout, x' denotes the transpose of a vector x , and B' denotes the transpose of a matrix B .

Definition 6. An m -player differential game of fixed prescribed duration $[0, T]$ is of the *affine-quadratic* type if $U_j = \mathbb{R}^{p_j}$ ($j \in M$) and

$$\begin{aligned} f(t, x, u) &= A(t)x + \sum_{i \in M} B_i(t)u_i + c(t), \\ g_j(t, x, u) &= -\frac{1}{2} \left(x' Q_j(t)x + \sum_{i \in M} u'_i R_j^i(t)u_i \right), \\ S_j(x) &= -\frac{1}{2} x'_j Q_j^f x, \end{aligned}$$

where $A(\cdot)$, $B_j(\cdot)$, $Q_j(\cdot)$, $R_j^i(\cdot)$ are matrices of appropriate dimensions, $c(\cdot)$ is an n -dimensional vector, all defined on $[0, T]$, and with continuous entries ($i, j \in M$). Furthermore, Q_j^f , $Q_j(\cdot)$ are symmetric, $R_j^j(\cdot) > 0$ ($j \in M$), and $R_j^j(\cdot) \geq 0$ ($i \neq j$, $i, j \in M$).

An affine-quadratic differential game is of the *linear-quadratic* type if $c \equiv 0$.

3.8.1 OLNE

For the affine-quadratic differential game formulated above, let us further assume that $Q^i(\cdot) \geq 0$, $Q_j^i \geq 0$. Then, under the open-loop information structure, player j 's payoff function $J_j([u_j, u_{-j}]; x^0, t^0 = 0)$, defined by (2.3), is a strictly concave function of $u_j(\cdot)$ for all permissible control functions $u_{-j}(\cdot)$ of the other players and for all $x^0 \in \mathbb{R}^n$. This then implies that the necessary conditions for OLNE derived in Sect. 3.3.1 are also sufficient, and every solution set of the first-order conditions provides an OLNE. Now, the Hamiltonian for player j is

$$H_j(x, u, \lambda_j, t) = -\frac{1}{2} \left(x' Q_j x + \sum_{i \in M} u'_i R_j^i u_i \right) + \lambda_j \left(Ax + c + \sum_{i \in M} B_i u_i \right),$$

whose maximization with respect to $u_j(t) \in \mathbb{R}^{p_j}$ yields the unique relation

$$u_j^*(t) = R_j^j(t)^{-1} B_j(t)' \lambda_j(t), \quad j \in M. \quad (\text{i})$$

Furthermore, the costate equations are

$$\dot{\lambda}_j = Q_j x^* - A' \lambda_j; \quad \lambda_j(T) = -Q_j^f x(T) \quad (j \in M), \quad (\text{ii})$$

and the optimal state trajectory is generated by

$$\dot{x}^* = Ax^* + c - \sum_{i \in M} B_i R_i^{i-1} B_i' \lambda_i; \quad x^*(0) = x^0. \quad (\text{iii})$$

This set of differential equations constitutes a two-point boundary value problem, the solution of which can be written, without any loss of generality, as $\{\lambda_j(t) = -K_j(t)x^*(t) - k_j(t), j \in M; x^*(t), t \in [0, T]\}$ where $K_j(\cdot)$ are $(n \times n)$ -dimensional matrices and $k_j(\cdot)$ are n -dimensional vectors. Now, substituting $\lambda_j = -K_j x^* - k_j$ ($j \in M$) into the costate equations (ii), we can arrive at the conclusion that K_j ($j \in M$) and k_j ($j \in M$) should then satisfy, respectively, the following two sets of matrix and vector differential equations:

$$\begin{aligned} \dot{K}_j + K_j A + A' K_j + Q_j^j - K_j \sum_{i \in M} B_i R_i^{i-1} B_i' K_i &= 0; \\ K_j(T) &= Q_j^f \quad (j \in M), \end{aligned} \quad (2.42)$$

and

$$\begin{aligned} \dot{k}_j + A' k_j + K_j c - K_j \sum_{i \in M} B_i R_i^{i-1} B_i' k_i &= 0; \\ k_j(T) &= 0 \quad (j \in M). \end{aligned} \quad (2.43)$$

The expressions for the OLNE strategies can then be obtained from (i) by substituting $\lambda_j = -K_j x^* - k_j$, and likewise the associated state trajectory for x^* follows from (iii).

The following theorem now captures this result (see, Başar and Olsder 1999, pp. 317–318).

Theorem 3. *For the m -player affine-quadratic differential game with $Q_j(\cdot) \geq 0$, $Q_j^f \geq 0$ ($j \in M$), let there exist a unique solution set $\{K_j, j \in M\}$ to the coupled set of matrix Riccati differential equations (2.42). Then, the differential game admits a unique OLNE solution given by*

$$\gamma_j^*(x^0, t) \equiv u_j^*(t) = -R_i^i(t)^{-1} B_i'(t) [K_j(t)x^*(t) + k_j(t)] \quad (j \in M),$$

where $\{k_j(\cdot), j \in M\}$ solve uniquely the set of linear differential equations (2.43) and $x^*(\cdot)$ denotes the corresponding OLNE state trajectory, generated by (iii), which can be written as

$$\begin{aligned}
x^*(t) &= \Phi(t, 0)x_0 + \int_0^t \Phi(t, \sigma)\eta(\sigma) d\sigma, \\
\frac{d}{dt}\Phi(t, \sigma) &= F(t)\Phi(t, \sigma); \quad \Phi(\sigma, \sigma) = I, \\
F(t) &:= A - \sum_{i \in M} B_i R_i^{i-1} B_i' K_i(t), \\
\eta(t) &:= c(t) - \sum_{i \in M} B_i R_i^{i-1} B_i' k_i(t).
\end{aligned}$$

Remark 6 (Nonexistence and multiplicity of OLNE). Note that the existence of OLNE for the affine-quadratic differential game hinges on the existence of a solution to the set of coupled matrix Riccati equations (2.42), since the second differential equation (2.43) always admits a solution, being linear in k_i 's. Further, the OLNE is unique, whenever the matrix solution to (2.42) is unique. It is within the realm of possibility, however, that an OLNE may not exist, just as a Nash equilibrium may not exist in static quadratic games (reaction planes may not have a common point of intersection) or there may be multiple OLNEs (using the earlier analogy to static games, reaction planes may have more than one point of intersection). Note also that even for the LQDG (i.e., when $c \equiv 0$), when we have $k_j \equiv 0$, $j \in M$, still the same possibilities (of nonexistence or multiplicity of OLNE) are valid.

An important point to note regarding the OLNE in Theorem 3 above is that the solution does not depend on all the parameters that define the affine-quadratic differential game, particularly the matrices $\{R_j^i, i \neq j, i, j = 1, 2\}$. Hence, the OLNE would be the same if g_j were replaced by

$$\tilde{g}_j(t, x, u_j) = -\frac{1}{2} \left(x' Q_j(t)x + u_j' R_j^j(t)u_j \right).$$

This is in fact not surprising in view of our earlier discussion in Sect. 3.7 on strategic equivalence. Under open-loop information structure, adding to $g_j(t, x, u)$ of one game any function of u_{-j} generates another game that is strategically equivalent to the first one and hence has the same set of OLNE strategies, and in this particular case, adding the term $(1/2) \sum_{i \neq j} u_i' R_j^i(t)u_i$ to g_j generates \tilde{g}_j . We can now go a step further, and subtract the term $(1/2) \sum_{i \neq j} u_i' R_i^i(t)u_i$ from \tilde{g}_j , and assuming also that the state weighting matrices $Q_j(\cdot)$ and Q_j^f are the same across all players (i.e., $Q(\cdot)$ and Q^f , respectively), we arrive at a single-objective function for all players (where we suppress dependence on t in the weighting matrices):

$$J_j(u(\cdot); x^0) =: J(u(\cdot); x^0) = -\frac{1}{2} \left(\int_0^T (x' Q x + \sum_{i \in M} u_i' R_i^i u_i) dt + x(T)' Q^f x(T) \right). \quad (2.44)$$

Hence, the affine-quadratic differential game where $Q_j(\cdot)$ and Q_j^f are the same across all players is strategically equivalent to a team problem, which, being deterministic, is in fact an optimal control problem. Letting $u := (u_1', \dots, u_m')'$,

$B := (B_1, \dots, B_m)$, and $R = \text{diag}(R_1^1, \dots, R_m^m)$, this affine-quadratic optimal control problem has state dynamics

$$\dot{x} = A(t)x + B(t)u(t) + c(t), \quad x(0) = x^0,$$

and payoff function

$$J(u(\cdot); x^0) = -\frac{1}{2} \left(\int_0^T (x' Q(t)x + u' R(t)u) dt + x(T)' Q^f x(T) \right),$$

where $R(\cdot) > 0$. Being strictly concave (and affine-quadratic), this optimal control problem admits a unique globally optimal solution, given by⁷

$$u^*(t) = -R(t)^{-1} B(t)' [K(t)x^*(t) + k(t)], \quad t \geq 0,$$

where $K(\cdot)$ is the unique nonnegative-definite solution of the matrix Riccati equation:

$$\dot{K} + KA + A'K + Q - KBR^{-1}B'K = 0, \quad K(T) = Q^f, \quad (2.45)$$

$k(\cdot)$ uniquely solves

$$\dot{k} + A'k + Kc - KBR^{-1}B'k = 0, \quad k(T) = 0 \quad (2.46)$$

and $x^*(\cdot)$ is generated by (3), with

$$F(t) = A - BR^{-1}B'K(t), \quad \eta(t) = c(t) - BR^{-1}B'k(t), \quad t \geq 0.$$

Note that for each block component of u , the optimal control can be written as

$$\gamma_j^*(x^0, t) = u_j^*(t) = -R_j^j(t)^{-1} B_j(t)' [K(t)x^*(t) + k(t)], \quad t \geq 0, j \in M, \quad (2.47)$$

which by strategic equivalence is the unique OLNE. The following corollary to Theorem 3 summarizes this result.

Corollary 1. *The special class of affine-quadratic differential games with open-loop information structure, where in Definition 6, $Q_j = Q \geq 0 \forall j \in M$ and $Q_j^f = Q^f \geq 0 \forall j \in M$, is strategically equivalent to a strictly concave optimal control problem and admits a unique OLNE, given by (2.47), where $K(\cdot)$ is the unique nonnegative-definite solution of (2.45), $k(\cdot)$ uniquely solves (2.46), and $x^*(\cdot)$ is the unique OLNE state trajectory as defined above.*

⁷This is a standard result in optimal control, which can be found in any standard text, such as Bryson et al. (1975).

Remark 7 (Strategic equivalence and symmetry). A special class of affine-quadratic differential games which fits into the framework covered by Corollary 1 is the class of *symmetric differential games*, where the players are indistinguishable (with B_j, Q_j, Q_j^f, R_j^j being the same across all players, that is, index j free). Hence, symmetric affine-quadratic differential games, with $Q_j = Q \geq 0 \forall j \in M$, $Q_j^f = Q^f \geq 0 \forall j \in M$, $R_j^j = \bar{R} > 0 \forall j \in M$, and $B_j = \bar{B} \forall j \in M$, admit a unique OLNE:

$$\gamma_j^*(x^0, t) = u_j^*(t) = -\bar{R}(t)^{-1} \bar{B}(t)' [K(t)x^*(t) + k(t)], \quad t \geq 0, j \in M,$$

where $K(\cdot)$ and $k(\cdot)$ uniquely solve

$$\dot{K} + KA + A'K + Q - mK\bar{B}\bar{R}^{-1}\bar{B}'K = 0, \quad K(T) = Q^f, \quad (2.48)$$

and

$$\dot{k} + A'k + Kc - mK\bar{B}\bar{R}^{-1}\bar{B}'k = 0, \quad k(T) = 0,$$

and $x^*(\cdot)$ is as defined before.

Remark 8 (Zero-sum differential games). A special class of nonzero-sum differential games is zero-sum differential games, where in the general framework, $m = 2$ and $J_2 \equiv -J_1 =: J$. The (two) players in this case have totally opposing objectives, and hence what one would be minimizing, the other one would be maximizing. Nash equilibrium in this case corresponds to the *saddle-point equilibrium*, and if (γ_1^*, γ_2^*) is one such pair of strategies, with player 1 as minimizer (of J) and player 2 as maximizer, they satisfy the pair of saddle-point inequalities:

$$\bar{J}(\gamma_1^*, \gamma_2) \leq \bar{J}(\gamma_1^*, \gamma_2^*) \leq \bar{J}(\gamma_1, \gamma_2^*), \quad \forall \gamma_j \in \Gamma_j, j = 1, 2. \quad (2.49)$$

Affine-quadratic zero-sum differential games are defined as in Definition 6, with $m = 2$ and (suppressing dependence on the time variable t)

$$\begin{aligned} g_2 \equiv -g_1 &=: g(x, u_1, u_2, t) \\ &= \frac{1}{2}(x'Qx + u_1'R_1u_1 - u_2'R_2u_2), \quad Q \geq 0, R_i > 0, i = 1, 2, \end{aligned}$$

$$S_2(x) \equiv -S_1(x) =: S(x) = \frac{1}{2}x'Q^f x, \quad Q^f \geq 0.$$

Note, however, that this formulation cannot be viewed as a special case of two-player affine-quadratic nonzero-sum differential games with nonpositive-definite weighting on the states and negative-definite weighting on the controls of the players in their payoff functions (which makes the payoff functions strictly concave

in individual players' controls – making their individual maximization problems automatically well defined), because here maximizing player (player 2 in this case) has nonnegative-definite weighting on the state, which brings up the possibility of player 2's optimization problem to be unbounded. To make the game well defined, we have to assure that it is convex-concave. Convexity of J in u_1 is readily satisfied, but for concavity in u_2 , we have to impose an additional condition. It turns out that (see Başar and Bernhard 1995; Başar and Olsder 1999) a practical way of checking strict concavity of J in u_2 is to assure that the following matrix Riccati differential equation has a continuously differentiable nonnegative-definite solution over the interval $[0, T]$, that is, there are no conjugate points:

$$\dot{\hat{S}} + \hat{S}A + A'\hat{S} + Q + \hat{S}B_2R_2^{-1}B_2'\hat{S} = 0, \quad \hat{S}(T) = Q^f. \quad (2.50)$$

Then, one can show that the game admits a unique saddle-point solution in open-loop policies, which can be obtained directly from Theorem 3 by noticing that $K_2 = -K_1 =: \hat{K}$ and $k_2 = -k_1 =: \hat{k}$, which satisfy

$$\dot{\hat{K}} + \hat{K}A + A'\hat{K} + Q - \hat{K}(B_1R_1^{-1}B_1' - B_2R_2^{-1}B_2')\hat{K} = 0, \quad \hat{K}(T) = Q^f, \quad (2.51)$$

and

$$\dot{\hat{k}} + A'\hat{k} + \hat{K}c - \hat{K}(B_1R_1^{-1}B_1' - B_2R_2^{-1}B_2')\hat{k} = 0, \quad \hat{k}(T) = 0. \quad (2.52)$$

Under the condition of existence of a well-defined solution to (2.50), the matrix Riccati differential equation (2.51) admits a unique continuously differentiable nonnegative-definite solution, and the open-loop saddle-point (OLSP) strategies for the players, satisfying (2.49), are given by

$$\begin{aligned} \gamma_1^*(x^0, t) &= -R_1^{-1}B_1'[\hat{K}(t)x^*(t) + \hat{k}(t)], \\ \gamma_2^*(x^0, t) &= R_2^{-1}B_2'[\hat{K}(t)x^*(t) + \hat{k}(t)], \quad t \geq 0, \end{aligned}$$

where $x^*(\cdot)$ is the saddle-point state trajectory, generated by

$$\dot{x} = (A - (B_1'R_1^{-1}B_1' - B_2R_2^{-1}B_2')\hat{K})x - (B_1'R_1^{-1}B_1 - B_2R_2^{-1}B_2)\hat{k} + c, \quad x(0) = x^0.$$

Hence, the existence of an OLSP hinges on the existence of a nonnegative-definite solution to the matrix Riccati differential equation (2.50), which as indicated is related to the nonexistence of conjugate points in the interval $[0, T]$,⁸ which in turn is related to whether the game in the infinite-dimensional function space (Hilbert

⁸The existence of a conjugate point in $[0, T]$ implies that there exists a sequence of policies by the maximizer which can drive the value of the game arbitrarily large, that is, the upper value of the game is infinite.

space in this case) is convex-concave or not, as mentioned earlier. For details, we refer to Başar and Bernhard (1995).

3.8.2 SFNE

We now turn to affine-quadratic differential games (cf. Definition 6) with state-feedback information structure. We have seen earlier (cf. Theorem 2) that the CLNE strategies can be obtained from the solution of coupled HJB partial differential equations. For affine-quadratic differential games, these equations can be solved explicitly, since their solutions admit a general quadratic (in x) structure, as we will see shortly. This also readily leads to a set of SFNE strategies which are expressible in closed form. The result is captured in the following theorem, which follows from Theorem 2 by using in the coupled HJB equations the structural specification of the affine-quadratic game (cf. Definition 6), testing the solution structure $V_j(t, x) = -\frac{1}{2}x'Z_j(t)x - x'\zeta_j(t) - n_j(t)$, $j \in M$, showing consistency, and equating like powers of x to arrive at differential equations for Z_j , ζ_j , and n_j (see, Başar and Olsder 1999, pp. 323–324).

Theorem 4. *For the m -player affine-quadratic differential game introduced in Definition 6, with $Q_j(\cdot) \geq 0$, $Q_j^f \geq 0$ ($j \in M$), let there exist a set of matrix-valued functions $Z_j(\cdot) \geq 0$, $j \in M$, satisfying the following m -coupled matrix Riccati differential equations:*

$$\begin{aligned} \dot{Z}_j + Z_j \tilde{F} + \tilde{F}' Z_j + \sum_{i \in M} Z_i B_i R_i^{i-1} R_j^i R_i^{i-1} B_i' Z_i + Q_j &= 0; \\ Z_j(T) &= Q_j^f, \end{aligned} \quad (2.53)$$

where

$$\tilde{F}(t) := A(t) - \sum_{i \in M} B_i(t) R_i^i(t)^{-1} B_i(t)' Z_i(t). \quad (2.54)$$

Then, under the state-feedback information structure, the differential game admits an SFNE solution, affine in the current value of the state, given by

$$\gamma_j^*(x, t) = -R_j^j(t)^{-1} B_j(t)' [Z_j(t)x(t) + \zeta_j(t)], \quad j \in M, \quad (2.55)$$

where ζ_j ($j \in M$) are obtained as the unique solution of the coupled linear differential equations

$$\dot{\zeta}_j + \tilde{F}' \zeta_j + \sum_{i \in M} Z_i B_i R_i^{i-1} R_j^i R_i^{i-1} B_i' \zeta_i + Z_j \beta = 0; \quad \zeta_j(T) = 0, \quad (2.56)$$

with

$$\beta := c - \sum_{i \in M} B_i R_i^{i-1} B_i' \zeta_i. \quad (2.57)$$

The corresponding values of the payoff functionals are

$$\bar{J}_j^* = V_j(x^0, 0) = -\frac{1}{2} x^{0'} Z_j(0) x^0 - x^{0'} \zeta_j(0) - n_j(0), \quad j \in M, \quad (2.58)$$

where $n_j(\cdot)$ ($j \in M$) are obtained as unique continuously differentiable solutions of

$$\dot{n}_j + \beta' \zeta_j + \frac{1}{2} \sum_{i \in M} \zeta_i' B_i R_i^{i-1} R_j^i R_i^{i-1} B_i' \zeta_i = 0; \quad n_j(T) = 0. \quad (2.59)$$

Remark 9. Note that the “nonnegative-definiteness” requirement imposed on $Z_j(\cdot)$ is a consequence of the fact that $V_j(x, t) \geq 0 \forall x \in \mathbb{R}^n, t \in [0, T]$, this latter feature being due to the eigenvalue restrictions imposed a priori on $Q_j(\cdot)$, Q_j^f , and $R_j^i(\cdot)$, $i, j \in M$. Finally, the corresponding “Nash” values for the payoff functionals follow from the fact that $V_j(x, t)$ is the value function for player j at SFNE, at any point (x, t) . We also note that Theorem 4 provides only one set of SFNE strategies for the affine-quadratic game under consideration, and it does not attribute any uniqueness feature to this solution set. What can be shown, however, is the uniqueness of SFNE when the players are restricted at the outset to affine memoryless state-feedback strategies (Başar and Olsder 1999).

Remark 10. The result above extends readily to more general affine-quadratic differential games where the payoff functions of the players contain additional terms that are linear in x , that is, with g_j and S_j in Definition 6 extended, respectively, to

$$g_j = -\frac{1}{2} \left(x' [Q_j(t)x + 2l_j(t)] + \sum_{i \in M} u_i' R_i^j u_i \right); \quad S_j(x) = -\frac{1}{2} x' [Q_j^f x + 2l_j^f],$$

where $l_j(\cdot)$ is a known n -dimensional vector-valued function, continuous on $[0, T]$, and l_j^f is a fixed n -dimensional vector, for each $j \in M$. Then, the statement of Theorem 4 remains intact, with only the differential equation (2.56) that generates $\zeta_j(\cdot)$ now reading:

$$\dot{\zeta}_j + \tilde{F}' \zeta_j + \sum_{i \in M} Z_i B_i R_i^{i-1} R_j^i R_i^{i-1} B_i' \zeta_i + Z_j \beta + l_j = 0; \quad \zeta_j(T) = l_j^f.$$

When comparing SFNE with the OLNE, one question that comes up is whether there is the counterpart of Corollary 1 in the case of SFNE. The answer is no, because adding additional terms to g_j that involve controls of other players

generally leads to a different optimization problem faced by player j , since u_i 's for $i \neq j$ depend on x and through it on u_j . Hence, in general, a differential game with state-feedback information structure cannot be made strategically equivalent to a team (and hence optimal control) problem. One can, however, address the issue of simplification of the set of sufficient conditions (particularly the coupled matrix Riccati differential equations (2.53)) when the game is symmetric. Let us use the same setting as in Remark 7, but also introducing a common notation \hat{R} for the weighting matrices $R_j^i, i \neq j$ for the controls of the other players appearing in player j 's payoff function, and for all $j \in M$ (note that this was not an issue in the case of open-loop information since R_j^i 's, $i \neq j$ were not relevant to the OLNE), and focusing on symmetric SFNE, we can now rewrite the SFNE strategy (2.55) for player j as

$$\gamma_j^*(x, t) = -\bar{R}(t)^{-1}\bar{B}(t)'[Z(t)x(t) + \zeta(t)], \quad j \in M, \quad (2.60)$$

where $Z(\cdot) \geq 0$ solves

$$\dot{Z} + Z\tilde{F} + \tilde{F}'Z + Z\bar{B}\bar{R}^{-1}\bar{B}'Z + \sum_{i \neq j, i \in M} Z\bar{B}\bar{R}^{-1}\hat{R}\bar{R}^{-1}\bar{B}'^f, \quad (2.61)$$

with (from (2.54))

$$\tilde{F} := A(t) - m\bar{B}(t)\bar{R}(t)^{-1}\bar{B}(t)'Z(t). \quad (2.62)$$

Substituting this expression for \tilde{F} into (2.61), we arrive at the following alternative (more revealing) representation:

$$\begin{aligned} \dot{Z} + ZA + A'Z - Z\bar{B}\bar{R}^{-1}\bar{B}'Z + Q - (m-1)Z\bar{B}\bar{R}^{-1}[2\bar{R} - \hat{R}]\bar{R}^{-1}\bar{B}'Z \\ = 0; \quad Z(T) = Q^f. \end{aligned} \quad (2.63)$$

Using the resemblance to the matrix Riccati differential equation that arises in standard optimal control (compare it with the differential equation (2.48) for K in Remark 7), we can conclude that (2.63) admits a unique continuously differentiable nonnegative-definite solution whenever the condition

$$2\bar{R} - \hat{R} > 0, \quad (2.64)$$

holds. This condition can be interpreted as players placing relatively more weight on their self-controls (in their payoff functions) than on each of the other individual players. In fact, if the weights are equal, then $\bar{R} = \hat{R}$, and (2.63) becomes equivalent to (2.48); this is of course not surprising since a symmetric game with $\bar{R} = \hat{R}$ is essentially an optimal control problem (players have identical payoff functions), for which OL and SF solutions have the same underlying Riccati differential equations.

Now, to complete the characterization of the SFNE for the symmetric differential game, we have to write down the differential equation for ζ_j in Theorem 4, that is (2.56), using the specifications imposed by symmetry. Naturally, it now becomes independent of the index j and can be simplified to the form below:

$$\dot{\zeta} + \left[A' - Z\bar{B}\bar{R}^{-1}[(2m-1)\bar{R} - (m-1)\hat{R}]\bar{R}^{-1}\bar{B}' \right] \zeta + Zc = 0; \zeta(T) = 0. \quad (2.65)$$

The following corollary to Theorem 4 now captures the main points of the discussion above.

Corollary 2. *For the symmetric affine-quadratic differential game introduced above, let the matrix Riccati differential equation (2.63) admit a unique continuously differentiable nonnegative-definite solution $Z(\cdot)$. Then the game admits a CLNE solution, which is symmetric across all players, and given by*

$$\gamma_j^*(x, t) = -\bar{R}(t)^{-1}\bar{B}(t)'[Z(t)x(t) + \zeta(t)], \quad t \geq 0, j \in M,$$

where $\zeta(\cdot)$ is generated uniquely by (2.65). If, furthermore, the condition (2.64) holds, then $Z(\cdot)$ exists and is unique.

Remark 11 (Zero-sum differential games with SF information). The counterpart of Remark 8 on saddle-point equilibrium can also be derived under state-feedback information structure, this time specializing Theorem 4 to the two-player zero-sum differential game. Using the same setting as in Remark 8, it follows by inspection from Theorem 4 that $Z_2 = -Z_1 =: \hat{Z}$ and $\zeta_2 = \zeta_1 =: \hat{\zeta}$, where the differential equations satisfied by \hat{Z} and $\hat{\zeta}$ are precisely the ones satisfied by \hat{K} and \hat{k} in the OL case, that is, (2.51) and (2.52), respectively. Under the condition of the existence of well-defined (unique continuously differentiable nonnegative-definite) solution to the matrix Riccati differential equation (2.51), the state-feedback saddle-point (SFSP) strategies for the players, satisfying (2.49), are given by (directly from Theorem 4)

$$\gamma_1^*(x, t) = -R_1^{-1}B_1'[\hat{K}(t)x(t) + \hat{k}(t)], \quad \gamma_2^*(x, t) = R_2^{-1}B_2'[\hat{K}(t)x(t) + \hat{k}(t)], \quad t \geq 0.$$

Note that these are in the same form as the OLSP strategies, with the difference being that they are now functions of the *actual* current value of the state instead of the *computed* value (as in the OLSP case). Another difference between the OLSP and SFSP is that the latter does not require an a priori concavity condition to be imposed, and hence whether there exists a solution to (2.50) is irrelevant under state-feedback information; this condition is replaced by the existence of a solution to (2.51), which is less restrictive Başar and Bernhard (1995). Finally, since the forms of the OLSP and SFSP strategies are the same, they generate the same state trajectory (and hence lead to the same value for J), provided that the corresponding existence conditions are satisfied.

4 Stackelberg Equilibria

In the previous sections, the assumption was that the players select their strategies simultaneously, without any communication. Consider now a different scenario: a two-player game where one player, the leader, makes her decision before the other player, the follower.⁹ Such a sequence of moves was first introduced by von Stackelberg in the context of a duopoly output game; see, von Stackelberg (1934).

Denote by L the leader and by F the follower. Suppose that $u_L(t)$ and $u_F(t)$ are, respectively, the control vectors of L and F . The control constraints $u_L(t) \in U_L$ and $u_F(t) \in U_F$ must be satisfied for all t . The state dynamics and the payoff functionals are given as before by (2.1), (2.2), and (2.3), where we take the initial time to be $t^0 = 0$, without any loss of generality. As with the Nash equilibrium, we will define an open-loop Stackelberg equilibrium (OLSE). We will also introduce what is called feedback (or Markovian)-Stackelberg equilibrium (FSE), which uses state feedback information and provides the leader only time-incremental lead advantage.

4.1 Open-Loop Stackelberg Equilibria (OLSE)

When both players use open-loop strategies, μ_L and μ_F , their control paths are determined by $u_L(t) = \mu_L(x^0, t)$ and $u_F(t) = \mu_F(x^0, t)$, respectively. Here μ_j denotes the open-loop strategy of player j .

The game proceeds as follows. At time $t = 0$, the leader announces her control path $u_L(\cdot)$ for $t \in [0, T]$. Suppose, for the moment, that the follower believes in this announcement. The best she can do is then to select her own control path $u_F(\cdot)$ to maximize the objective functional

$$J_F = \int_0^T g_F(x(t), u_L(t), u_F(t), t) dt + S_F(x(T)), \quad (2.66)$$

subject to the state dynamics

$$\dot{x}(t) = f(x(t), u_L(t), u_F(t), t) \quad x(0) = x^0, \quad (2.67)$$

and the control constraint

$$u_F(t) \in U_F. \quad (2.68)$$

⁹The setup can be easily extended to the case of several followers. A standard assumption is then that the followers play a (Nash) simultaneous-move game vis-a-vis each other, and a sequential game vis-a-vis the leader (Başar and Olsder 1999).

This is a standard optimal control problem. To solve it, introduce the follower's Hamiltonian

$$H_F(x(t), \lambda_F(t), u_F(t), u_L(t), t) \\ = g_F(x(t), u_F(t), u_L(t), t) + \lambda_F f(x(t), u_F(t), u_L(t), t),$$

where the adjoint variable $\lambda_F = \lambda_F(t)$ is an n -vector. Suppose that the Hamiltonian H_F is strictly concave in $u_F \in U_F$, where U_F is a convex set. Then the maximization of H_F with respect to u_F , for $t \in [0, T]$, uniquely determines $u_F(t)$ as a function of t, x, u_L , and λ_F , which we write as

$$u_F(t) = R(x(t), t, u_L(t), \lambda_F(t)). \quad (2.69)$$

This defines the follower's *best reply (response)* to the leader's announced time path $u_L(\cdot)$.

The follower's costate equations and their boundary conditions in this maximization problem are given by

$$\dot{\lambda}_F(t) = -\frac{\partial}{\partial x} H_F, \\ \lambda_F(T) = \frac{\partial}{\partial x} S_j(x(T)).$$

Substituting the best response function R into the state and costate equations yields a two-point boundary-value problem. The solution of this problem, $(x(t), \lambda_F(t))$, can be inserted into the function R . This represents the follower's optimal behavior, given the leader's announced time path $u_L(\cdot)$.

The leader can replicate the follower's arguments. This means that, since she knows everything the follower does, the leader can calculate the follower's best reply R to any $u_L(\cdot)$ that she may announce. The leader's problem is then to select a control path $u_L(\cdot)$ that maximizes her payoff given F 's response, that is, maximization of

$$J_L = \int_0^T g_L(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t)), t) dt + S_L(x(T)), \quad (2.70)$$

subject to

$$\dot{x}(t) = f(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t)), t), \quad x(0) = x^0, \\ \dot{\lambda}_F(t) = -\frac{\partial}{\partial x} H_F(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t))), t), \\ \lambda_F(T) = \frac{\partial}{\partial x} S_F(x(T)),$$

and the control constraint

$$u_L(t) \in U_L.$$

Note that the leader's dynamics include two state equations, one governing the evolution of the original state variables x and a second one accounting for the evolution of λ_F , the adjoint variables of the follower, which are now treated as state variables. Again, we have an optimal control problem that can be solved using the maximum principle. To do so, we introduce the leader's Hamiltonian

$$\begin{aligned} H_L(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t)), \lambda_L(t), \theta(t)) \\ = g_L(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t)), t) \\ + \lambda_L(t) f(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t)), t) \\ + \theta(t) \left(-\frac{\partial}{\partial x} H_F(x(t), \lambda_F(t), R(x(t), t, u_L(t), \lambda_F(t)), u_L(t), t) \right), \end{aligned}$$

where $\lambda_L = \lambda_L(t)$ is the n -vector of costate variables appended to the state equation for $x(t)$, with the boundary conditions

$$\lambda_L(T) = \frac{\partial}{\partial x} S_L(x(T)),$$

and $\theta = \theta(t)$ is the vector of n costate variables appended to the state equation for $\lambda_F(t)$, satisfying the initial condition

$$\theta(0) = 0.$$

This initial condition is a consequence of the fact that $\lambda_F(0)$ is “free,” i.e., unrestricted, being free of any soft constraint in the payoff function, as opposed to $x(T)$ which enters a terminal reward term. The following theorem now collects all this for the OLSE (see, Başar and Olsder 1999, pp. 409–410).

Theorem 5. *For the two-player open-loop Stackelberg differential game formulated in this subsection, let $u_L^*(t) = \mu_L^*(x^0, t)$ be the leader's open-loop equilibrium strategy and $u_F^*(t) = \mu_F^*(x^0, t)$ be the follower's. Let the solution to the follower's optimization problem of maximizing J_F given by (2.66) subject to the state equation (2.67) and control constraint (2.68) exist and be uniquely given by (2.69). Then,*

- (i) *The leader's open-loop Stackelberg strategy μ_L^* maximizes (2.70) subject to the given control constraint and the $2n$ -dimensional differential equation system for x and λ_F (given after (2.70)) with mixed boundary specifications.*

(ii) *The follower's open-loop Stackelberg strategy μ_F^* is (2.69) with u_L replaced by u_L^* .*

Remark 12. In the light of the discussion in this subsection leading to Theorem 5, it is possible to write down a set of necessary conditions (based on the maximum principle) which can be used to solve for L 's open-loop strategy μ_L^* . Note that, as mentioned before, in this maximization problem in addition to the standard state (differential) equation with specified initial conditions, we also have the costate differential equation with specified terminal conditions, and hence the dynamic constraint for the maximization problem involves a $2n$ -dimensional differential equation with mixed boundary conditions (see the equations for x and λ_F following (2.70)). The associated Hamiltonian is then H_L , defined prior to Theorem 5, which has as its arguments two adjoint variables, λ_L and θ , corresponding to the differential equation evolutions for x and λ_F , respectively. Hence, from the maximum principle, these new adjoint variables satisfy the differential equations:

$$\begin{aligned}\dot{\lambda}_L(t) &= -\frac{\partial}{\partial x} H_L(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t)), \lambda_L(t), \theta(t), t), \\ \lambda_L(T) &= \frac{\partial}{\partial x} S_L(x(T)), \\ \dot{\theta}(t) &= -\frac{\partial}{\partial \lambda_F} H_L(x(t), u_L(t), R(x(t), t, u_L(t), \lambda_F(t)), \lambda_L(t), \theta(t), t), \\ \theta(0) &= 0.\end{aligned}$$

Finally, $u_L^*(t) = \mu_L^*(x^0, t)$ is obtained from the maximization of the Hamiltonian H_L (where we suppress dependence on t):

$$u_L^* = \arg \max_{u_L \in U_L} H_L(x, u_L, R(x, u_L, \lambda_F), \lambda_L, \theta).$$

4.2 Feedback Stackelberg Equilibria (FSE)

We now endow both players with state-feedback information, as was done in the case of SFNE, which is a memoryless information structure, not allowing the players to recall even the initial value of the state, x^0 , except at $t = 0$. In the case of Nash equilibrium, this led to a meaningful solution, which also had the appealing feature of being subgame perfect and strongly time consistent. We will see in this subsection that this appealing feature does not carry over to Stackelberg equilibrium when the leader announces her strategy in advance for the entire duration of the game, and in fact the differential game becomes ill posed. This will force us to introduce, again under the state-feedback information structure, a different concept of Stackelberg equilibrium, called *feedback Stackelberg*, where the strong time

consistency is imposed at the outset. This will then lead to a derivation that parallels the one for SFNE.

Let us first address “ill-posedness” of the classical Stackelberg solution when the players use state-feedback information, in which case their strategies are mappings from $\mathbb{R}^n \times [0, T]$ where the state-time pair (x, t) maps into U_L and U_F , for L and F , respectively. Let us denote these strategies by $\gamma_L \in \Gamma_L$ and $\gamma_F \in \Gamma_F$, respectively. Hence, the realizations of these strategies lead to the control actions (or control paths): $u_L(t) = \gamma_L(x, t)$ and $u_F(t) = \gamma_F(x, t)$, for L and F , respectively. Now, in line with the OLSE we discussed in the previous subsection, under the Stackelberg equilibrium, the leader L announces at time zero her strategy $\gamma(x, t)$ and commits to using this strategy throughout the duration of the game. Then the follower F reacts rationally to L 's announcement, by maximizing her payoff function. Anticipating this, the leader selects a strategy that maximizes her payoff functional subject to the constraint imposed by the best response of F .

First let us look at the follower's optimal control problem. Using the dynamic programming approach, we have the Hamilton-Jacobi-Bellman (HJB) equation characterizing F 's best response to an announced $\gamma_L \in \Gamma_L$:

$$-\frac{\partial}{\partial t} V_F(x, t) = \max_{u_F \in U_F} \left\{ g_F(x, u_F(t), \gamma_L(x, t), t) + \frac{\partial}{\partial x} V_F(x, t) f(x(t), u_F(t), \gamma_L(x, t), t) \right\},$$

where V_F is the value function of F , which has the terminal condition $V_F(x, T) = S_F(x(T))$. Note that, for each fixed $\gamma_L \in \Gamma_L$, the maximizing control for F on the RHS of the HJB equation above is a function of the current time and state and hence is an element of Γ_F . Thus, F 's maximization problem and its solution are compatible with the state-feedback information structure, and hence we have a well-defined problem at this stage. The dependence of this best response on γ_L , however, will be quite complex (much more than in the open-loop case), since what we have is a functional dependence in an infinite-dimensional space. Nevertheless, at least formally, we can write down this relationship as a best reaction function, $\tilde{R} : \Gamma_L \rightarrow \Gamma_F$, for the follower:

$$\gamma_F = \tilde{R}(\gamma_L). \quad (2.71)$$

Now, L can make this computation too, and according to the Stackelberg equilibrium concept, which is also called *global Stackelberg solution* (see, Başar and Olsder 1999), she has to maximize her payoff under the constraints imposed by this reaction function and the state dynamics that is formally

$$\max_{\gamma_L \in \Gamma_L} J_L(\gamma_L, \tilde{R}(\gamma_L))$$

Leaving aside the complexity of this optimization problem (which is not an optimal control problem of the standard type because of the presence of the reaction function which depends on the entire strategy of L over the full-time interval of the game), we note that this optimization problem is ill-posed since for each choice of $\gamma_L \in \Gamma_L$, $J_L(\gamma_L, \hat{R}(\gamma_L))$ is not a real number but generally a function of the initial state x^0 , which is not available to L ; hence, what we have is a multi-objective optimization problem, and not a single-objective one, which makes the differential game with the standard (global) Stackelberg equilibrium concept ill-posed. One way around this difficulty would be to allow the leader (as well as the follower) recall the initial state (and hence modify their information sets to $v(x(t), x^0, t)$) or even have full memory on the state (in which case, v is $v(x(s), s \leq t; t)$), which would make the game well-posed, but requiring a different set of tools to obtain the solution (see, e.g., Başar and Olsder 1980; Başar and Selbuz 1979 and Chap. 7 of Başar and Olsder 1999), which also has connections to incentive designs and inducement of collusive behavior, further discussed in the next section of this chapter. We should also note that including x^0 in the information set also makes it possible to obtain the global Stackelberg equilibrium under mixed information sets, with F 's information being inferior to that of L , such as $v_L(x(t), x^0, t)$ for L and $v(x^0, t)$ for F . Such a differential game would also be well-defined.

Another way to resolve the ill-posedness of the global Stackelberg solution under state-feedback information structure is to give the leader only a *stagewise* (in the discrete-time context) first-mover advantage; in continuous time, this translates into an instantaneous advantage at each time t (Başar and Haurie 1984). This pointwise (in time) advantage leads to what is called a *feedback Stackelberg equilibrium (FSE)*, which is also strongly time consistent (Başar and Olsder 1999). The characterization of such an equilibrium for $j \in \{L, F\}$ involves the HJB equations

$$\left\{ -\frac{\partial}{\partial t} V_j(x, t) \right\}_{j=1,2} = \text{Sta} \left\{ g_j(x, [u_F, u_L]) + \frac{\partial}{\partial x} V_j(x, t) f(x, [u_F, u_L], t) \right\}_{j=1,2}, \quad (2.72)$$

where the ‘‘Sta’’ operator on the RHS solves, for each (x, t) , for the Stackelberg equilibrium solution of the static two-player game in braces, with player 1 as leader and player 2 as follower. More precisely, the pointwise (in time) best response of F to $\gamma_L \in \Gamma_L$ is

$$\hat{R}(x, t; \gamma_L(x, t)) = \arg \max_{u_F \in U_F} \left\{ g_F(x, [u_F, \gamma_L(x, t)], t) + \frac{\partial}{\partial x} V_F(x, t) f(x, [u_F, \gamma_L(x, t)], t) \right\},$$

and taking this into account, L solves, again pointwise in time, the maximization problem:

$$\max_{u_L \in U_L} \left\{ g_L(x), ([\hat{R}(x, t; u_L), u_L], t) + \frac{\partial}{\partial x} V_L(x, t) f(x, [\hat{R}(x, t; u_L), u_L], t) \right\}.$$

Denoting the solution to this maximization problem by $u_L = \hat{\gamma}_L(x, t)$, an FSE for the game is then the pair of state-feedback strategies:

$$\left(\hat{\gamma}_L(x, t), \hat{\gamma}_F(x, t) = \hat{R}(x, t; \hat{\gamma}_L(x, t)) \right) \quad (2.73)$$

Of course, following the lines we have outlined above, it should be obvious that explicit derivation of this pair of strategies depends on the construction of the value functions, V_L and V_F , satisfying the HJB equations (2.72). Hence, to complete the solution, one has to solve (2.72) for $V_L(x, t)$ and $V_F(x, t)$ and use these functions in (2.73). The main difficulty here is, of course, in obtaining explicit solutions to the HJB equations, which however can be done in some classes of games, such as those with linear dynamics and quadratic payoff functions (in which case V_L and V_F will be quadratic in x) (Başar and Olsder 1999). We provide some evidence of this solvability through numerical examples in the next subsection.

4.3 An Example: Construction of Stackelberg Equilibria

Consider the example of Sect. 3.7 but now with player 1 as the leader (from now on referred to as player L) and player 2 as the follower (player F). Recall that player j 's optimization problem and the underlying state dynamics are

$$\max_{u_j} \left\{ J_j = \int_0^\infty e^{-\rho t} \left(u_j(t) \left(\kappa - \frac{1}{2} u_j(t) \right) - \frac{1}{2} \varphi x^2(t) \right) dt \right\}, \quad j=L, F, \quad (2.74)$$

$$\dot{x}(t) = u_L(t) + u_F(t) - \alpha x(t), \quad x(0) = x^0, \quad (2.75)$$

where φ and κ are positive parameters and $0 < \alpha < 1$. We again suppress the time argument henceforth when no ambiguity may arise. We discuss below both OLSE and FSE, but with horizon length infinite. This will give us an opportunity to introduce, in this context, also the infinite-horizon Stackelberg differential game.

4.3.1 Open-Loop Stackelberg Equilibrium (OLSE).

To obtain the best reply of the follower to the leader's announcement of the path $u_L(t)$, we introduce the Hamiltonian of player F :

$$H_F(x, u_L, u_F) = u_F \left(\kappa - \frac{1}{2} u_F \right) - \frac{1}{2} \varphi x^2 + q_F (u_L + u_F - \alpha x),$$

where q_F is the follower's costate variable associated with the state variable x . H_F being quadratic and strictly concave in u_F , it has a unique maximum:

$$u_F = \kappa + q_F, \quad (2.76)$$

where (from the maximum principle) q_F satisfies

$$\dot{q}_F = \rho q_F - \frac{\partial}{\partial x} H_F = (\rho + \alpha)q_F + \varphi x, \quad \lim_{t \rightarrow \infty} e^{-\rho t} q_F(t) = 0, \quad (2.77)$$

and with (2.76) used in the state equation, we have

$$\dot{x}(t) = u_L(t) + \kappa + q_F(t) - \alpha x(t), \quad x(0) = x^0. \quad (2.78)$$

Now, one approach here would be first to solve the two differential equations (2.77) and (2.78) and next to substitute the solutions in (2.76) to arrive at follower's best reply, i.e., $u_F(t) = R(x(t), u_L(t), q_F(t))$. Another approach would be to postpone the resolution of these differential equations and instead use them as dynamic constraints in the leader's optimization problem:

$$\begin{aligned} \max_{u_L} \left\{ J_L = \int_0^\infty e^{-\rho t} \left(u_L \left(\kappa - \frac{1}{2} u_L \right) - \frac{1}{2} \varphi x^2 \right) dt \right\} \\ \dot{q}_F = (\rho + \alpha)q_F + \varphi x, \quad \lim_{t \rightarrow \infty} e^{-\rho t} q_F(t) = 0, \\ \dot{x} = u_L + \kappa + q_F - \alpha x, \quad x(0) = x^0. \end{aligned}$$

This is an optimal control problem with two state variables (q_F and x) and one control variable (u_L). Introduce the leader's Hamiltonian:

$$\begin{aligned} H_L(x, u_L, q_F, q_L, \theta) = u_L \left(\kappa - \frac{1}{2} u_L \right) - \frac{1}{2} \varphi x^2 + \theta ((\rho + \alpha)q_F + \varphi x) \\ + q_L (u_L + \kappa + q_F - \alpha x), \end{aligned}$$

where θ and q_L are adjoint variables associated with the two state equations in the leader's optimization problem. Being quadratic and strictly concave in u_L , H_L also admits a unique maximum, given by

$$u_L = \kappa + q_L,$$

and we have the state and adjoint equations:

$$\dot{\theta} = \rho \theta - \frac{\partial}{\partial q_F} H_L = -\theta \alpha - q_L, \quad \theta(0) = 0,$$

$$\begin{aligned}\dot{q}_L &= \rho q_L - \frac{\partial}{\partial x} H_L = (\rho + \alpha)q_L + \varphi(x - \theta), & \lim_{t \rightarrow \infty} e^{-\rho t} q_L(t) &= 0, \\ \dot{x} &= \frac{\partial}{\partial q_L} H_L = u_L + \kappa + q_F - \alpha x, & x(0) &= x^0, \\ \dot{q}_F &= \frac{\partial}{\partial \theta} H_L = (\rho + \alpha)q_F + \varphi x, & \lim_{t \rightarrow \infty} e^{-\rho t} q_F(t) &= 0.\end{aligned}$$

Substituting the expression for u_L in the differential equation for x , we obtain a system of four differential equations, written in matrix form as follows:

$$\begin{pmatrix} \dot{\theta} \\ \dot{q}_L \\ \dot{x} \\ \dot{q}_F \end{pmatrix} = \begin{pmatrix} -\alpha & -1 & 0 & 0 \\ -\varphi & \rho + \alpha & \varphi & 0 \\ 0 & 1 & -\alpha & 1 \\ 0 & 0 & \varphi & \rho + \alpha \end{pmatrix} \begin{pmatrix} \theta \\ q_L \\ x \\ q_F \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2\kappa \\ 0 \end{pmatrix}.$$

Solving the above system yields (θ, q_L, x, q_F) . The last step would be to insert the solutions for q_F and q_L in the equilibrium conditions

$$u_F = \kappa + q_F, \quad u_L = \kappa + q_L,$$

to obtain the open-loop Stackelberg equilibrium controls u_L and u_F .

4.3.2 Feedback-Stackelberg Equilibrium (FSE).

To obtain the FSE, we first have to consider the infinite-horizon version of (2.72) and compute the best response of F to $u_L = \gamma_L(x)$. The maximization problem faced by F has the associated steady-state HJB equation for the current-value function $\mathcal{V}_F(x)$ (with the value function defined as $V_F(x, t) = e^{-\rho t} \mathcal{V}_F(x)$):

$$\rho \mathcal{V}_F(x) = \max_{u_F} \left[u_F \left(\kappa - \frac{1}{2} u_F \right) - \frac{1}{2} \varphi x^2 + \frac{\partial}{\partial x} \mathcal{V}_F(x) (u_L + u_F - \alpha x) \right]. \quad (2.79)$$

Maximization of the RHS yields (uniquely, because of strict concavity)

$$u_F = \kappa + \frac{\partial}{\partial x} \mathcal{V}_F(x). \quad (2.80)$$

Note that the above reaction function of the follower does not directly depend on the leader's control u_L , but only indirectly, through the state variable.

Accounting for the follower's response, the leader's HJB equation is

$$\rho \mathcal{V}_L(x) = \max_{u_L \geq 0} \left[u_L \left(\kappa - \frac{1}{2} u_L \right) - \frac{1}{2} \varphi x^2 + \frac{\partial}{\partial x} \mathcal{V}_L(x) \left(u_L + \kappa + \frac{\partial}{\partial x} \mathcal{V}_F(x) - \alpha x \right) \right], \quad (2.81)$$

where $\mathcal{V}_L(x)$ denotes the leader's current-value function. Maximizing the RHS yields

$$u_L = \kappa + \frac{\partial}{\partial x} \mathcal{V}_L(x).$$

Substituting in (2.81) leads to

$$\begin{aligned} \rho \mathcal{V}_L(x) &= \left(\kappa + \frac{\partial}{\partial x} \mathcal{V}_L(x) \right) \left(\kappa - \frac{1}{2} \left(\kappa + \frac{\partial}{\partial x} \mathcal{V}_L(x) \right) \right) \\ &\quad - \frac{1}{2} \varphi x^2 + \frac{\partial}{\partial x} \mathcal{V}_L(x) \left(\frac{\partial}{\partial x} \mathcal{V}_L(x) + \frac{\partial}{\partial x} \mathcal{V}_F(x) + 2\kappa - \alpha x \right). \end{aligned} \quad (2.82)$$

As the game at hand is of the linear-quadratic type, we can take the current value functions to be general quadratic. Accordingly, let

$$\mathcal{V}_L(x) = \frac{A_L}{2} x^2 + B_L x + C_L, \quad (2.83)$$

$$\mathcal{V}_F(x) = \frac{A_F}{2} x^2 + B_F x + C_F, \quad (2.84)$$

be, respectively, the leader's and the follower's current-value functions, where the six coefficients are yet to be determined. Substituting these structural forms in (2.82) yields

$$\begin{aligned} \rho \left(\frac{A_L}{2} x^2 + B_L x + C_L \right) &= \frac{1}{2} (A_L^2 - \varphi + 2(A_F - \alpha) A_L) x^2 \\ &\quad + (A_L (B_L + B_F + 2\kappa) + (A_F - \alpha) B_L) x + \frac{1}{2} (\kappa^2 + B_L^2) + (B_F + 2\kappa) B_L. \end{aligned}$$

Using (2.79), (2.80), and (2.83)–(2.84), we arrive at the following algebraic equation for the follower:

$$\begin{aligned} \rho \left(\frac{A_F}{2} x^2 + B_F x + C_F \right) &= \frac{1}{2} (A_F^2 - \varphi + 2(A_L - \alpha) A_F) x^2 \\ &\quad + (A_F (B_F + B_L + 2\kappa) + (A_L - \alpha) B_F) x + \frac{1}{2} (\kappa^2 + B_F^2) + (B_L + 2\kappa) B_F. \end{aligned}$$

By comparing the coefficients of like powers of x , we arrive at the following six-equation, nonlinear algebraic system:

$$\begin{aligned} 0 &= A_L^2 + (2A_F - 2\alpha - \rho) A_L - \varphi, \\ 0 &= A_L (B_L + B_F + 2\kappa) + (A_F - \alpha - \rho) B_L, \end{aligned}$$

$$\begin{aligned}
0 &= \frac{1}{2} (\kappa^2 + B_L^2) + (B_F + 2\kappa) B_L - \rho C_L, \\
0 &= A_F^2 - \varphi + (2A_L - 2\alpha - \rho) A_F, \\
0 &= A_F (B_F + B_L + 2\kappa) + (A_L - \alpha - \rho) B_F, \\
0 &= \frac{1}{2} (\kappa^2 + B_F^2) + (B_L + 2\kappa) B_F - \rho C_F.
\end{aligned}$$

The above system generally admits multiple solutions. One can eliminate some of these based on, e.g., convergence to an asymptotically globally stable steady state. Let the sextuple $(A_L^S, B_L^S, C_L^S, A_F^S, B_F^S, C_F^S)$ denote a solution to the above system, satisfying the additional desirable properties. Then, a pair of FSE strategies is given by

$$\begin{aligned}
u_F &= \kappa + \mathcal{V}'_F(x) = A_F^S x + B_F^S, \\
u_L &= \kappa + \mathcal{V}'_L(x) = A_L^S x + B_L^S.
\end{aligned}$$

4.4 Time Consistency of Stackelberg Equilibria

When, at an initial instant of time, the leader announces a strategy she will use throughout the game, her goal is to influence the follower's strategy choice in a way that will be beneficial to her. *Time consistency* addresses the following question: given the option to re-optimize at a later time, will the leader stick to her original plan, i.e., the announced strategy and the resulting time path for her control variable? If it is in her best interest to deviate, then the leader will do so, and the equilibrium is then said to be *time inconsistent*. An inherently related question is then why would the follower, who is a rational player, believe in the announcement made by the leader at the initial time if it is not credible? The answer is clearly that she would not.

In most of the Stackelberg differential games, it turns out that the OLSE is time inconsistent, that is, the leader's announced control path $u_L(\cdot)$ is not credible. Markovian or feedback Stackelberg equilibria (SFE), on the other hand, are subgame perfect and hence time consistent; they are in fact *strongly time consistent*, which refers to the situation where the restriction of leader's originally announced strategy to a shorter time interval (sharing the same terminal time) is still SFE and regardless of what evolution the game had up to the start of that shorter interval.

The OLSE in Sect. 4.3 is time inconsistent. To see this, suppose that the leader has the option of revising her plan at time $\tau > 0$ and to choose a new decision rule $u_L(\cdot)$ for the remaining time span $[\tau, \infty)$. Then she will select a rule that satisfies $\theta(\tau) = 0$ (because this choice will fulfill the initial condition on the costate θ). It can be shown, by using the four state and costate equations $(\dot{x}, \dot{q}_F, \dot{q}_L, \dot{\theta})$, that for some instant of time, $\tau > 0$, it will hold that $\theta(\tau) \neq 0$. Therefore, the leader

will want to announce a new strategy at time τ , and this makes the original strategy time inconsistent, i.e., the new strategy does not coincide with the restriction of the original strategy to the interval $[\tau, \infty)$.

Before concluding this subsection, we make two useful observations.

Remark 13. Time consistency (and even stronger, strong time consistency) of FSE relies on the underlying assumption that the information structure is state feedback and hence without memory, that is, at any time t , the players do **not** remember the history of the game.

Remark 14. In spite of being time inconsistent, the OLSE can still be a useful solution concept for some short-term horizon problems, where it makes sense to assume that the leader will not be tempted to re-optimize at an intermediate instant of time.

5 Memory Strategies and Collusive Equilibria

5.1 Implementing Memory Strategies in Differential Games

As mentioned earlier, by memory strategies we mean that the players can, at any instant of time, recall any specific past information. The motivation for using memory strategies in differential games is in reaching through an equilibrium a desirable outcome that is not obtainable noncooperatively using open-loop or state-feedback strategies. Loosely speaking, this requires that the players *agree* (implicitly, or without taking on any binding agreement) on a desired trajectory to follow throughout the game (typically a cooperative solution) and are willing to implement a punishment strategy if a deviation is observed. Richness of an information structure, brought about through incorporation of memory, enables such monitoring.

If one party realizes, or remembers, that, in the past, the other party deviated from an agreed-upon strategy, it implements some pre-calculated punishment. Out of the fear of punishment, the players adhere to the Pareto-efficient path, which would be unobtainable in a strictly noncooperative game.

A punishment is conceptually and practically attractive only if it is *effective*, i.e., it deprives a player of the benefits of a defection, and *credible*, i.e., it is in the best interest of the player(s) who did not defect to implement this punishment. In this section, we first introduce the concept of non-Markovian strategies and the resulting Nash equilibrium and next illustrate these concepts through a simple example.

Consider a two-player infinite-horizon differential game, with state equation

$$\dot{x}(t) = f(x(t), u_1(t), u_2(t), t), \quad x(0) = x^0.$$

To a pair of controls $(u_1(t), u_2(t))$, there corresponds a unique trajectory $x(\cdot)$ emanating from x^0 . Player j 's payoff is given by

$$J_j(u_1(t), u_2(t); x^0) = \int_0^\infty e^{-\rho_j t} g_j(x(t), u_1(t), u_2(t), t) dt, \quad j = 1, 2,$$

where $g_j(x(t), u_1(t), u_2(t), t)$ is taken to be bounded and continuously differentiable.¹⁰ As before, the control set of player j is U_j and the state set X is identical to \mathbb{R}^n .

Heretofore, a strategy has been defined as a mapping from player's information space to her control set. Unfortunately, this direct approach poses formidable mathematical difficulties in the present context; therefore, we will define a strategy as an infinite sequence of approximate constructions, called δ -strategies. For player j , consider the sequence of times $t_j = i\delta, i = 0, 1, \dots$, where δ is a fixed positive number. For any time interval $[t_j, t_{i+1})$, let \mathcal{U}_j^i be the set of measurable control functions $u_{j,i} : [t_j, t_{i+1}) \rightarrow U_j$, and let $\mathcal{U}^i = \mathcal{U}_1^i \times \mathcal{U}_2^i$. A δ -strategy for player j is a sequence $\Delta_j^\delta = (\Delta_{j,i})_{i=0,1,\dots}$ of mappings

$$\begin{aligned} \Delta_{j,0} &\in \mathcal{U}_j^0, \\ \Delta_{j,i} &= U^0 \times U^1 \times \dots \times U^{i-1} \rightarrow \mathcal{U}_j^i \quad \text{for } j = 1, 2, \dots \end{aligned}$$

A strategy for player j is an infinite sequence of δ -strategies:

$$\Delta_j = \left\{ \Delta_j^{\delta_n} : \delta_n \rightarrow 0, n = 1, 2, \dots \right\}.$$

Note that this definition implies that the information set of player j at time t is

$$\{(u_1(s), u_2(s)), 0 \leq s < t\},$$

that is, the entire control history up to (but not including) time t . So when players choose δ -strategies, they are using, at successive sample times t_j , the accumulated information to generate a pair of measurable controls $(u_1^\delta(\cdot), u_2^\delta(\cdot))$ which, in turn, generate a unique trajectory $x^\delta(\cdot)$ and thus, a unique outcome $w^\delta = (w_1^\delta, w_2^\delta) \in \mathbb{R}^2$, where $\bar{\delta} = (\delta, \delta')$, and

$$w_j^\delta = \int_0^\infty e^{-\rho_j t} g_j(x^\delta(t), u_1^\delta(t), u_2^\delta(t), t) dt.$$

¹⁰This assumption allows us to use the strong-optimality concept and avoid introducing additional technicalities.

An *outcome* of the strategy pair $\bar{\Delta}$ is a pair $\bar{w} \in \mathbb{R}^2$, which is a *limit* of the sequence $\{w^{\bar{\delta}_n}\}$ of the outcomes of δ -strategy pairs $\bar{\Delta}^{\bar{\delta}_n} = (\Delta_1^{\bar{\delta}_n}, \Delta_2^{\bar{\delta}_n})$ when n tends to infinity. With a strategy pair, the initial state and time are thus associated with a set $v(t^0, x^0; \bar{\Delta})$ of possible outcomes. (Note that we have used the obvious extension to a non-zero initial time t^0 .) The game is well defined if, for any strategy pair $\bar{\Delta}$ and any initial conditions (t^0, x^0) , the set of outcomes $v(t^0, x^0; \bar{\Delta})$ is nonempty.¹¹

Definition 7. A strategy pair $\bar{\Delta}^*$ is a Nash equilibrium at (t^0, x^0) if, and only if,

1. the outcome set $v(t^0, x^0; \bar{\Delta}^*)$ reduces to a singleton $w^* = (w_1^*, w_2^*)$;
2. for all strategy pairs $\bar{\Delta}^{(1)} \triangleq (\Delta_1, \Delta_2^*)$ and $\bar{\Delta}^{(2)} \triangleq (\Delta_1^*, \Delta_2)$, the following holds for $j = 1, 2$:

$$(w_1, w_2) \in v(t^0, x^0; \bar{\Delta}^{(i)}) \Rightarrow w_j \leq w_j^*.$$

The equilibrium condition for the strategy pair is valid only at (t^0, x^0) . This implies, in general, that the Nash equilibrium that was just defined is not subgame perfect.

Definition 8. A strategy pair $\bar{\Delta}^*$ is a subgame-perfect Nash equilibrium at (t^0, x^0) if, and only if,

1. given a control pair $\bar{u}(\cdot) : [t^0, t] \rightarrow U_1 \times U_2$ and the state $x(t)$ reached at time t , we define the prolongation of $\bar{\Delta}^*$ at $(t, x(t))$ as $\{\Delta^{*\delta_n} : \delta_n \rightarrow 0, n = 1, 2, \dots\}$ defined by

$$\begin{aligned} & \Delta^{*\delta_n} (\bar{u}_{[t, t+\delta_n]}, \dots, \bar{u}_{[t+i\delta_n, t+(i+1)\delta_n]}) \\ &= \Delta^{*\delta_n} (\bar{u}_{[0, \delta_n]}, \bar{u}_{[\delta_n, 2\delta_n]}, \dots, \bar{u}_{[t+i\delta_n, t+(i+1)\delta_n]}); \end{aligned}$$

2. the prolongation of $\bar{\Delta}^*$ at $(t, x(t))$ is again an equilibrium at $(t, x(t))$.

Before providing an illustrative example, we make a couple of points in the following remark.

Remark 15. 1. The information set was defined here as the entire control history. An alternative definition is $\{x(s), 0 \leq s < t\}$, that is, each player bases her decision on the entire past state trajectory. Clearly, this definition requires less

¹¹Here, and in the balance of this section, we depart from our earlier convention of state-time ordering (x, t) , and use the reverse ordering (t, x) .

memory capacity and hence may be an attractive option, particularly when the differential game involves more than two players. (See Tolwinski et al. 1986 for details.)

2. The consideration of memory strategies in differential games can be traced back to Varaiya and Lin (1963), Friedman (1971), and Krassovski and Subbotin (1977). Their setting was (mainly) zero-sum differential games, and they used memory strategies as a convenient tool for proving the existence of a solution. Başar used memory strategies in the 1970s to show how richness of and redundancy in information structures could lead to *informationally nonunique* Nash equilibria (Başar 1974, 1975, 1976, 1977) and how the richness and redundancy can be exploited to solve for global Stackelberg equilibria (Başar 1979, 1982; Başar and Olsder 1980; Başar and Selbuz 1979) and to obtain incentive designs (Başar 1985). The exposition above follows Tolwinski et al. (1986) and Haurie and Pohjola (1987), where the setting is nonzero-sum differential games and the focus is on the construction of cooperative equilibria.

5.2 An Example

Consider a two-player differential game where the evolution of the state is described by

$$\dot{x}(t) = (1 - u_1(t)) u_2(t), \quad x(0) = x^0 > 0, \quad (2.85)$$

where $0 < u_j(t) < 1$. The players maximize the following objective functionals:

$$J_1(u_1(t), u_2(t); x^0) = \alpha \int_0^\infty e^{-\rho t} (\ln u_1(t) + x(t)) dt,$$

$$J_2(u_1(t), u_2(t); x^0) = (1 - \alpha) \int_0^\infty e^{-\rho t} (\ln(1 - u_1(t))(1 - u_2(t)) + x(t)) dt,$$

where $0 < \alpha < 1$ and $0 < \rho \leq 1/4$.

Suppose that the two players wish to implement a cooperative solution noncooperatively by using non-Markovian strategies and threats.

Step 1: Determine Cooperative Outcomes. Assume that these outcomes are given by the joint maximization of the sum of players' payoffs. To solve this optimal control problem, we introduce the current-value Hamiltonian (we suppress the time argument):

$$\mathcal{H}(u_1, u_2, x, \lambda) = \alpha \ln u_1 + (1 - \alpha) \ln(1 - u_1)(1 - u_2) + x + q(1 - u_1)u_2,$$

where q is the current-value adjoint variable associated with the state equation (2.85). Necessary and sufficient optimality conditions are

$$\begin{aligned}\dot{x} &= (1 - u_1) u_2, & x(0) &= x^0 > 0, \\ \dot{q} &= \rho q - 1, & \lim_{t \rightarrow \infty} e^{-\rho t} q(t) &= 0, \\ \frac{\partial \mathcal{H}}{\partial u_1} &= \frac{\alpha}{u_1} - \frac{(1 - \alpha)}{(1 - u_1)} - q u_2 = 0, \\ \frac{\partial \mathcal{H}}{\partial u_2} &= -\frac{(1 - \alpha)}{(1 - u_2)} - q(1 - u_1) = 0.\end{aligned}$$

It is easy to verify that the unique optimal solution is given by

$$\begin{aligned}(u_1^*, u_2^*) &= \left(\alpha \rho, \frac{1 - \rho}{1 - \alpha \rho} \right), & x^*(t) &= x^0 + (1 - \rho)t, \\ J_1(u_1^*(\cdot), u_2^*(\cdot); x^0) &= \frac{\alpha}{\rho} \left(\ln \alpha \rho + x^0 + \frac{1 - \rho}{\rho} \right), \\ J_2(u_1^*(\cdot), u_2^*(\cdot); x^0) &= \frac{1 - \alpha}{\rho} \left(\ln(1 - \alpha) \rho + x^0 + \frac{1 - \rho}{\rho} \right).\end{aligned}$$

Note that both optimal controls satisfy the constraints $0 < u_j(t) < 1$, $j = 1, 2$.

Step 2: Compute Nash-Equilibrium Outcomes. As the game is of the linear-state variety,¹² open-loop and state-feedback Nash equilibria coincide. We therefore proceed with the derivation of the OLNE, which is easier to solve. To determine this equilibrium, we first write the players' current-value Hamiltonians:

$$\begin{aligned}\mathcal{H}_1(u_1, u_2, x, q_1) &= \alpha (\ln u_1 + x) + q_1 (1 - u_1) u_2, \\ \mathcal{H}_2(u_1, u_2, x, q_2) &= (1 - \alpha) (\ln(1 - u_1) (1 - u_2) + x) + q_2 (1 - u_1) u_2,\end{aligned}$$

where q_j is the costate variable attached by player j to the state equation (2.85). Necessary conditions for a Nash equilibrium are

$$\begin{aligned}\dot{x} &= (1 - u_1) u_2, & x(0) &= x^0 > 0, \\ \dot{q}_1 &= \rho q_1 - \alpha, & \lim_{t \rightarrow \infty} e^{-\rho t} q_1(t) &= 0, \\ \dot{q}_2 &= \rho q_2 - (1 - \alpha), & \lim_{t \rightarrow \infty} e^{-\rho t} q_2(t) &= 0, \\ \frac{\partial}{\partial u_1} \mathcal{H}_1 &= \frac{\alpha}{u_1} - q_1 u_2 = 0,\end{aligned}$$

¹²In a linear-state differential game, the objective functional, the salvage value and the dynamics are linear in the state variables. For such games, it holds that a feedback strategy is constant, i.e., independent of the state and hence open-loop and state-feedback Nash equilibria coincide.

$$\frac{\partial}{\partial u_2} \mathcal{H}_2 = -\frac{(1-\alpha)}{(1-u_2)} - q_2(1-u_1) = 0.$$

It is easy to check that the Nash equilibrium is unique and is given by

$$\begin{aligned} (\bar{u}_1, \bar{u}_2) &= \left(\frac{1-k}{2}, \frac{1+k}{2} \right), \quad \bar{x}(t) = x^0 + \left(\frac{1+k}{4} \right) t, \\ J_1(\bar{u}_1(\cdot), \bar{u}_2(\cdot); x^0) &= \frac{\alpha}{\rho} \left(\ln \left(\frac{1-k}{2} \right) + x^0 + \frac{1+k}{2\rho} - 1 \right), \\ J_2(\bar{u}_1(\cdot), \bar{u}_2(\cdot); x^0) &= \frac{1-\alpha}{\rho} \left(\ln \rho + x^0 + \frac{1+k}{2\rho} - 1 \right), \end{aligned}$$

where $k = \sqrt{1-4\rho}$. Note that the equilibrium controls satisfy the constraints $0 < u_j(t) < 1$, $j = 1, 2$, and as expected in view of the game structure, they are constant over time.

Step 3: Construct a Collusive Equilibrium. We have thus so far obtained

$$\begin{aligned} (w_1^*, w_2^*) &= (J_1(u_1^*(\cdot), u_2^*(\cdot); x^0), J_2(u_1^*(\cdot), u_2^*(\cdot); x^0)), \\ (\bar{w}_1, \bar{w}_2) &= (J_1(\bar{u}_1(\cdot), \bar{u}_2(\cdot); x^0), J_2(\bar{u}_1(\cdot), \bar{u}_2(\cdot); x^0)). \end{aligned}$$

Computing the differences

$$\begin{aligned} w_1^* - \bar{w}_1 &= \frac{\alpha}{\rho} \left(\ln \left(\frac{1-k}{2\rho} \right) + \frac{1+\rho^2-3\rho+k}{2\rho} \right), \\ w_2^* - \bar{w}_2 &= \frac{1-\alpha}{\rho} \left(\ln(1-\alpha) + \frac{1-3\rho+k}{2\rho} \right), \end{aligned}$$

we note that they are independent of the initial state x^0 and that their signs depend on the parameter values. For instance, if we have the following restriction on the parameter values:

$$0 < \alpha < \min \left(\frac{1-k}{2\rho \exp\left(\frac{3\rho-1-\rho^2-k}{2\rho}\right)}, 1 - \exp\left(\frac{3\rho-1+k}{2\rho}\right) \right),$$

then $w_1^* > \bar{w}_1$ and $w_2^* > \bar{w}_2$. Suppose that this is true. What remains to be shown is then that by combining the cooperative (Pareto-optimal) controls with the state-feedback (equivalent to open-loop, in this case) Nash strategy pair,

$$(\gamma_1(x), \gamma_2(x)) = (\bar{u}_1, \bar{u}_2) = \left(\frac{1-k}{2}, \frac{1+k}{2} \right),$$

we can construct a subgame-perfect equilibrium strategy in the sense of Definition 8.

Consider a strategy pair

$$\bar{\Delta}_j = \left\{ \bar{\Delta}_j^{*\delta_n} : \delta_n \rightarrow 0, n = 1, 2, \dots \right\},$$

where, for $j = 1, 2$, $\bar{\Delta}_j^{*\delta}$ is defined as follows:

$$\Delta_j^{*\delta} = (\Delta_{j,i})_{i=0,1,2,\dots},$$

with

$$\Delta_{j,0}^* = u_{j,0}^*(\cdot),$$

$$\Delta_{j,i}^* = \begin{cases} u_{j,i}^*(\cdot), & \text{if } \bar{u}(s) = \bar{u}^*(s) \text{ for almost } s \leq i\delta, \\ \varphi_j(x(j\delta)) = \bar{u}_j, & \text{otherwise,} \end{cases}$$

for $i = 1, 2, \dots$, where $u_{j,i}^*(\cdot)$ denotes the restriction truncation of $u_j^*(\cdot)$ to the subinterval $[i\delta, (i+1)\delta]$, $i = 0, 1, 2, \dots$, and $x(i\delta)$ denotes the state observed at time $t = i\delta$.

The strategy just defined is known as a *trigger strategy*. A statement of the trigger strategy, as it would be made by a player, is “At time t , I implement my part of the optimal solution if the other player has never cheated up to now. If she cheats at t , then I will retaliate by playing the state-feedback Nash strategy from t onward.” It is easy to show that this trigger strategy constitutes a subgame-perfect equilibrium.

Remark 16. It is possible in this differential-game setting to define a retaliation period of finite length, following a deviation. Actually, the duration of this period can be designed to discourage any player from defecting. Also, in the above development and example, we assumed that a deviation is instantaneously detected. This may not necessarily be the case, and in such situations we can consider a detection lag. For an example of a trigger strategy with a finite retaliation period and detection lag, see Hämmäläinen et al. (1984).

6 Conclusion

This chapter has provided an overview of the theory of nonzero-sum differential games formulated in continuous time and without any stochastic elements. Only noncooperative aspects of the theory have been covered, primarily under two different solution concepts: Nash equilibrium and Stackelberg equilibrium and several of their variants. The importance of information structures in such dynamic games has been emphasized, with special focus on open-loop and state-feedback information structures. The additional degrees of freedom memory strategies bring in in inducing specific behavior on the part of the players has also been discussed,

and several special structures of differential games, such as linear-quadratic (or affine-quadratic) games, symmetric games, and zero-sum differential games, have also been covered, with some illustrative examples. The chapter has also emphasized the important role strategic equivalence plays in solvability of some classes of differential games.

There are several other issues very relevant to the topic and material of this chapter, which are covered by selected other chapters in the *Handbook*. These involve dynamic games described in discrete time, concave differential games with coupled state constraints defined over infinite horizon, dynamic games with an infinite number of players (more precisely, mean-field games), zero-sum differential games (with more in-depth analysis than the coverage in this chapter), games with stochastic elements (more precisely, stochastic games), mechanism designs, and computational methods, to list just a few.

References

- Başar T (1974) A counter example in linear-quadratic games: existence of non-linear Nash solutions. *J Optim Theory Appl* 14(4):425–430
- Başar T (1975) Nash strategies for M-person differential games with mixed information structures. *Automatica* 11(5):547–551
- Başar T (1976) On the uniqueness of the Nash solution in linear-quadratic differential games. *Int J Game Theory* 5(2/3):65–90
- Başar T (1977) Informationally nonunique equilibrium solutions in differential games. *SIAM J Control* 15(4):636–660
- Başar T (1979) Information structures and equilibria in dynamic games. In: Aoki M, Marzollo A (eds) *New trends in dynamic system theory and economics*. Academic, New York, pp 3–55
- Başar T (1982) A general theory for Stackelberg games with partial state information. *Large Scale Syst* 3(1):47–56
- Başar T (1985) Dynamic games and incentives. In: Bagchi A, Jongen HTh (eds) *Systems and optimization*. Volume 66 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, pp 1–13
- Başar T (1989) Time consistency and robustness of equilibria in noncooperative dynamic games. In: Van der Ploeg F, de Zeeuw A (eds) *Dynamic policy games in economics*, North Holland, Amsterdam/New York, pp 9–54
- Başar T, Bernhard P (1995) H^∞ -optimal control and related minimax design problems: a dynamic game approach, 2nd edn. Birkhäuser, Boston
- Başar T, Haurie A (1984) Feedback equilibria in differential games with structural and modal uncertainties. In: Cruz JB Jr (ed) *Advances in large scale systems*, Chapter 1. JAE Press Inc., Connecticut, pp 163–201
- Başar T, Olsder GJ (1980) Team-optimal closed-loop Stackelberg strategies in hierarchical control problems. *Automatica* 16(4):409–414
- Başar T, Olsder GJ (1999) *Dynamic noncooperative game theory*. SIAM series in *Classics in Applied Mathematics*. SIAM, Philadelphia
- Başar T, Selbuz H (1979) Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems. *IEEE Trans Autom Control* AC-24(2):166–179
- Bellman R (1957) *Dynamic programming*. Princeton University Press, Princeton
- Blaquière A, Gérard F, Leitmann G (1969) *Quantitative and qualitative games*. Academic, New York/London
- Bryson AE Jr, Ho YC (1975) *Applied optimal control*. Hemisphere, Washington, DC

- Carlson DA, Haurie A, Leizarowitz A (1991) Infinite horizon optimal control: deterministic and stochastic systems, vol 332. Springer, Berlin/New York
- Case J (1969) Toward a theory of many player differential games. *SIAM J Control* 7(2):179–197
- Dockner E, Jørgensen S, Long NV, Sorger G (2000) Differential games in economics and management science. Cambridge University Press, Cambridge
- Engwerda J (2005) Linear-quadratic dynamic optimization and differential games. Wiley, New York
- Friedman A (1971) Differential games. Wiley-Interscience, New York
- Hämäläinen RP, Kaitala V, Haurie A (1984) Bargaining on whales: a differential game model with Pareto optimal equilibria. *Oper Res Lett* 3(1):5–11
- Haurie A, Pohjola M (1987) Efficient equilibria in a differential game of capitalism. *J Econ Dyn Control* 11:65–78
- Haurie A, Krawczyk JB, Zaccour G (2012) Games and dynamic games. World scientific now series in business. World Scientific, Singapore/Hackensack
- Isaacs R (1965) Differential games. Wiley, New York
- Jørgensen S, Zaccour G (2004) Differential games in marketing. Kluwer, Boston
- Krassovski NN, Subbotin AI (1977) Jeux différentiels. Mir, Moscow
- Lee B, Markus L (1972) Foundations of optimal control theory. Wiley, New York
- Leitmann G (1974) Cooperative and non-cooperative many players differential games. Springer, New York
- Mehlmann A (1988) Applied differential games. Springer, New York
- Selten R (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. *Int J Game Theory* 4:25–55
- Sethi SP, Thompson GL (2006) Optimal control theory: applications to management science and economics. Springer, New York. First edition 1981
- Starr AW, Ho YC (1969) Nonzero-sum differential games, Part I. *J Optim Theory Appl* 3(3): 184–206
- Starr AW, Ho YC (1969) Nonzero-sum differential games, Part II. *J Optim Theory Appl* 3(4): 207–219
- Tolwinski B, Haurie A, Leitmann G (1986) Cooperative equilibria in differential games. *J Math Anal Appl* 119:182–202
- Varaiya P, Lin J (1963) Existence of saddle points in differential games. *SIAM J Control Optim* 7:142–157
- von Stackelberg H (1934) Marktform und Gleichgewicht. Springer, Vienna [An English translation appeared in 1952 entitled *The theory of the market economy*. Oxford University Press, Oxford]
- Yeung DWK, Petrosjan L (2005) Cooperative stochastic differential games. Springer, New York